(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
29 March 2001 (29.03.2001)

PCT

(10) International Publication Number
WO 01/21636 A1

(51) International Patent Classification[7]: C07H 21/02, 21/04, C12P 21/06, 21/04, C12N 15/09, 15/00, C07K 1/00, A61K 39/095

(21) International Application Number: PCT/US00/26117

(22) International Filing Date: 22 September 2000 (22.09.2000)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
60/155,621 22 September 1999 (22.09.1999) US

(71) Applicant: THE RESEARCH FOUNDATION OF STATE UNIVERSITY OF NEW YORK [US/US]; State University of New York at Buffalo, Suite 200 UB Commons, 520 Lee Entrance, Amherst, NY 14228 (US).

(72) Inventors: RUSSO, Thomas; 19 Chapel Woods, Amherst, NY 14221 (US). CARLINO, Ulrike; 18 Abbottsford Place, Buffalo, NY 14213 (US).

(74) Agents: KADLE, Ranjana et al.; Hodgson, Russ, Andrews, Woods & Goodyear, LLP, One M & T Plaza, Suite 2000, Buffalo, NY 14203-2391 (US).

(81) Designated States *(national)*: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CR, CU, CZ, DE, DK, DM, DZ, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, UZ, VN, YU, ZA, ZW.

(84) Designated States *(regional)*: ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

**Published:**
— *With international search report.*
— *Before the expiration of the time limit for amending the claims and to be republished in the event of receipt of amendments.*

*For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.*

WO 01/21636 A1

# (54) Title: IDENTIFICATION OF A VACCINE CANDIDATE FROM AN EXTRAINTESTINAL ISOLATE OF E. COLI

(57) Abstract: The present invention is directed to identifying a gene, *iroN*$_{ec}$, from an extraintestinal isolate of E. coli, CP9. This gene was identified from CP9 by screening a library of 527 mutant derivatives of CP9 with active TnphoA fusion in human urine. Two mutant derivative CP9.45 and CP9.82 possessed increased PhoA activity in urine due to the Tn*pho*A insertion into *iroN*$_{ec}$. The product of this gene is an extracytoplasmic protein of 725 amino acids and can be used as vaccine against extraintestinal *E. coli* infections.

# IDENTIFICATION OF A VACCINE CANDIDATE FROM AN EXTRAINTESTINAL ISOLATE OF E. COLI

This applications claims the priority of U.S. provisional application serial no. 60/155,621 filed on September 22, 1999.

## FIELD OF THE INVENTION

This invention relates generally to the field of vaccines for bacterial infections. More particularly, this invention identifies a gene that is expressed in extraintestinal isolates of Escherichia coli (*E.coli*), and can be used as an immunogen in vaccine formulations.

## BACKGROUND OF THE INVENTION

*E.coli* is part of the normal intestinal flora where it does not cause infections. However, infections can occur if the bacteria gain entrance to other tissues and organs. This group of *E. coli* strains has been designated herein as extraintestinal pathogenic *Escherichia coli* (ExPEC). Extraintestinal infections (EIs) due to *E. coli* are common in all age groups and can involve nearly any organ or anatomical site. Typical EIs include urinary tract infection (UTI), meningitis (mainly in neonates and following neurosurgery), diverse intra-abdominal infection, pneumonia (particularly in hospitalized and institutionalized patients), intra-vascular device infection, osteomyelitis, and soft tissue infection, which usually occurs in the setting of tissue compromise. Bacteremia can accompany infection at any of these sites.

Extraintestinal isolates of *E. coli* are responsible for the majority of urinary tract infections (UTI). Eighty to ninety percent of ambulatory UTI, 73% in individuals over 50, and 25% of nosocomial UTI are due to extraintestinal strains of E. coli. Uncomplicated urethritis or cystitis occurs most commonly. However,

more severe sequelae of UTI includes pyelonephritis, intrarenal and perinephric abscess, and bacterimia with or without septic shock. Thus, it is clear that despite effective antimicrobial therapy, UTI due to E. coli causes considerable morbidity and mortality.

Although *E. coli* is considered to be a community-acquired pathogen, it also is the most frequently isolated gram-negative bacillus in long-term care facilities and hospitals. Severe illness and death can occur in otherwise healthy hosts, but adverse outcomes are considerably more likely in the presence of comorbid disease and impaired host defenses.

The scope and magnitude of infection caused by extraintestinal strains of *E. coli* is as great as any invasive bacterial pathogen. In fact, perhaps because disease due to extraintestinal isolates of *E. coli* is so common, the virulence and morbidity of this organism is often overlooked. As a result, extraintestinal strains of *E. coli* continue to be low profile "silent killers" imposing a medical-economic strain on the health care system.

Currently, no effective vaccine is available against ExPEC. To date, efforts to identify specific vaccine candidates against ExPEC have concentrated on virulence traits such as capsule, LPS, and pili. These studies have demonstrated that antibodies directed against these structures confer protection against homologous strains *in vivo*. However, the inherent marked antigenic variability of these components may limit their utility as vaccine candidates. Thus, there is an ongoing need for identification of novel effective strategies for the prevention and treatment of ExPEC infections including UTI.

Summary Of the Invention

The present invention is directed to identifying a gene, *iroN$_{ec}$*, from an extraintestinal isolate of *E. coli*. This gene was identified by transposon mutagenesis. Tn*phoA* mutagenesis was performed on the wild type isolate of *E.coli*, CP9, and a library of 527 mutants with active Tn*pho*A fusions was generated and saved. This library was subsequently screened for mutant derivatives with increased PhoA activity in the presence of urine. By this method, the *iroN$_{ec}$* gene was identified which encodes a protein of 725 amino acids. This protein has an extracytoplasmic location. DNA homology data and data demonstrating that its transcription is iron repressed supports its function as a siderophore receptor. High stringency Southern hybridization identified DNA sequences homologous to *iroN$_{ec}$* in 80-93% of ExPEC strains.

This protein or antigenic epitopes thereof can be used for inducing an immune response against extraintestinal *E.coli* infections. Data is presented to show that IroN$_{ec}$ is strongly immunogenic (without adjuvant) in mice. Further, antibodies developed against IroN$_{ec}$ were observed to be protective in a mouse intra-peritoneal challenge model. Mice immunized with IroN$_{ec}$ had diminished mortality after intra-peritoneal challenge with the *E. coli* strain CP9. Additionally, there was diminished growth of the challenged strain in the liver and spleen in animals immunized with IroN$_{ec}$ compared to non-immunized controls.

Thus, it is an object of the present invention to identify a gene, designated herein as *iroN$_{ec}$* in the extraintestinal isolates of *E.coli* with increased expression in human urine.

It is another object of the present invention to provide a polynucleotide that encodes the protein, $IroN_{ec}$.

It is another object of the present invention to provide polynucleotides that hybridize, preferably under high stringency conditions, with a polynucleotide encoding $IroN_{ec}$.

It is anther object of the present invention to provide peptides that are encoded by *iroN*$_{ec}$ and other polynucleotides of the present invention.

It is another object of the invention to provide antigenic compositions comprising the peptides or antigenic fragments thereof for the treatment or prevention of extraintestinal infections caused by ExPEC.

BRIEF DESCRIPTION OF THE DRAWINGS

Figure 1 is a schematic diagram of the *iroN*$_{ec}$ and the flanking sequence.

Figure 2A is a representation of the SDS-PAGE analysis of $IroN_{ec}$ purification. Lane 12 represents molecular weight markers. The approximate molecular weight of each marker is indicated. Lane 11 is non-purified sample of the induced culture and lanes 1-10 are eluted $IroN_{ec}$.

Figure 2B is a representation of Western blot analysis of SDS-PAGE from Figure 2A. An antibody to the recombinant $IroN_{ec}$ was used for detection. Lane 12 represents molecular weight markers having the indicated approximate molecular weights.

DESCRIPTION OF THE PREFERRED EMBODIMENTS

The present invention is directed to a gene whose expression is increased in human body fluids such as

urine, blood and ascites. This gene, termed *iroN*$_{ec}$, is represented by the sequence of SEQ ID NO:1.

In one aspect of this invention are provided polynucleotides that are at least 90% identical (homologous) over their entire length to the polynucleotide of SEQ ID NO:1, preferably at least 95% identical, and still more preferably at least 97% identical to the sequence of SEQ ID NO:1, or complementary sequences thereof.

The invention further relates to polynucleotides that hybridize to the sequence of SEQ ID NO:1 under high stringency hybridization conditions. High stringency conditions as used herein means hybridization will occur only if there is at least 90% homology, preferably at least 95% homology, and even more preferably at least 97-99% homology between the sequences. An example of high stringency hybridization conditions is overnight incubation at 42°C in 50% formamide, 5 X SSC, 50 mM sodium phosphate (pH 7.6), 5X Denhardt's solution, 10% dextran sulfate, and 20 μg/ml denatured, sheared salmon sperm DNA. Washing is carried out in 0.1 X SSC at about 65°C. Details on the stringent hybridization conditions are well known to those skilled in the art (see Sambrook et al., 1989, Molecular Cloning: A laboratory Manual, Second Edition, Cold Spring Harbor, NY). Another example of high stringency hybridization condition is presented in Example 2.

The polynucleotides of the present invention can be incorporated into vectors which can be used in expression systems for the production of polypeptides. A variety of expression systems are available and known to those skilled in the art.

In another embodiment of the invention are provided peptides that are encoded by the polynucleotides of the present invention. These peptides or antigenic

fragments thereof, can be used for diagnostic and immunotherapeutic purposes. "Antigen fragments" are those that are part of the whole protein and are specifically recognized by certain antibodies that also recognize the whole protein and which will generate an immune response that reduces or prevents the incidence or symptoms of *E.coli* infections.

In the present invention, *iroN*$_{ec}$ was identified by transposon mutagenesis. Transposons or transposable elements are DNA segments that can move in the genome from one site to another. They can also be moved from one bacterium to another by plasmid or other vehicles to be inserted into the host genome. Transposons may be simple or complex. Simple transposons have insertion sequences and no genes other than those involved in their own transposition into the target sequence. Complex transposons contain the insertion elements bracketing additional genes that encode properties such as drug resistance, carbohydrate metabolism, light generation, ice nucleation, and other properties which can function as selectable or screenable markers for the entire transposable elements. Thus, transposons are suitable for the use of identifying and cloning of bacterial genes which are turned off or on in the presence of a particular environment.

The method used for the identification of this gene has been described in Russo et al. (1996, Mol. Microbiology, 22:217-229), the disclosure of which is incorporated herein by reference. A transposon, Tn*phoA*, which is a Tn5 derivative, was used in the present invention. On correct insertion into an open reading frame, this transposon will produce a fusion protein with bacterial alkaline phosphatase (PhoA). Since PhoA is only active when in the periplasm, Tn*phoA* identifies genes whose products are exported across the

cell membrane. When a PhoA-specific chromogenic detection reagent (for example, XP; 5-bromo-4-chloro-3-indolyl phosphate) is used, a blue colony is only obtained when *TnphoA* has inserted into a gene encoding a protein containing the N-terminal export leader peptide sequence.

The pathogen CP9 (Russo et al., 1993, *Mol. Microbiol.*, 9:357-364), an *E. coli* blood isolate cultured from a patient with sepsis, hospitalized at the National Institutes of Health, was used as a model pathogen for identification of genes with increased expression in urine. It is characterized by growth in 80% normal human serum, ß-hemolysis, no known antibiotic resistance, O4/K54/H5 serotype, P pilus (class I PapG adhesin), Prs pilus (class III PapG adhesin), type 1 pilus, possession of a 36.2-kb cryptic plasmid (pJEG) and an aerobactin minus genotype. By DNA dot-blot assays, it has also been determined to be *sfa*, *ompT*, *cnf1* and *drb* positive (Johnson et al., 1997, *Infect. Immun.*, 65:2153-2159; Russo et al., 1993, *Mol. Microbiol.*, 9:357-364). It is highly virulent in a mouse UTI infection model (Russo et al., 1996, *Infec. Immun.*, 64:2343-2348). Recent studies have established that CP9 is part of a widely disseminated group of uropathogens that are characterized, in part, by possessing group 3 capsules, the O4 specific antigen, and Class 1 and 3 Pap adhesins (Johnson et al., 1997, *Infect. Immun.*, 65:2153-2159; Johnson et al., 1997, *Infect. Immun.*, 65:2153-2159).

All strains of *E.coli* were maintained at -80°C in 50% Luria-Bertani (L-B) medium and 50% glycerol. L-B broth consisted of 5 grams yeast extract, 10 grams tryptone, 10 grams NaCl per liter (L). Incubations were performed at 37°C unless otherwise described. For

plates, 15 grams of Bacto-Agar (Difco Laboratories , Detroit, MI) were added per L and kanamycin (kan) (40 $\mu$g/mL) or ampicillin (200 $\mu$g/ml) (Amresco, Solon, OH) were added where appropriate. Urine agar plates were made as described, using pooled urine from 5 healthy donors that did not have a history of having a urinary tract infection (Russo et al., 1996, *Mol. Microbiol.*, 22:217-229).

**EXAMPLE 1**

This embodiment describes the identification, cloning, and characterization of *iroN*$_{ec}$.

To identify genes with increased expression in human urine *ex vivo* that coded for extracytoplasmically located gene products, random Tn*phoA* mutagenesis was performed on CP9. A Tn*phoA* mutant library, consisting of 527 CP9 derivatives containing active Tn*phoA* fusions had been previously constructed (Russo et al., 1993, *Mol. Microbiol.*, 9:357-364) and was used in the present study.

The Tn*phoA* mutant library was screened to identify genes that coded for extra-cytoplasmically located gene products which had increased expression in human urine *ex vivo*. Transposon mutants were intially plated on human urine agar plates containing kanamycin (40$\mu$g/ml) and 5-bromo-4-chloro-3-indolyl-phosphate-p-toluidine salt (XP), as the colorimetric substrate for an indication of alkaline phosphatase activity. Mutants that appeared to have increased expression via these qualitative screens (as manifested by the relative blue color of the colonies) were confirmed with quantitative assays as follows. For urine and L-B assays quantitative alkaline phosphatase assays were performed as previously described (Russo et al., 1996, supra, Russo et al.,

1993, supra). In brief, CP9 and each of its mutant derivatives were grown overnight in L-B broth and human urine. Preliminary experiments established that PhoA expression was the same for CP9 derivatives containing an active PhoA fusion in *iroN*$_{ec}$ (CP9.45, CP9.82) in log phase and stationary phase grown cells. Cells were then washed, permeabilized, and p-nitrophenyl phosphate was added for detection of alkaline phosphatase activity respectively. To control for both endogenous bacterial alkaline phosphatase activity and any activity from the growth medium (e.g. urine) that may persist despite washing, alkaline phosphatase activity from CP9 was subtracted from the measured activity of its mutant derivatives.

The screening of the Tn*phoA* library resulted in the identification of CP9.45 and CP9.82 (Table 1).

Table 1

*E. coli* strains and plasmids used in this study

| Strain/ Plasmid | Genotype or other relevant characteristics | Derivation |
|---|---|---|
| Strains | | |
| CP9 | 04/K54/H5 | clinical blood isolate |
| CP9.45 | *iroN*$_{ec}$:Tn*phoA*, active Tn*phoA* fusion | Kan$^R$exconjugant of CP9/pRT291 x MM294/pPH1J1 |
| CP45 | *iroN*$_{ec}$:Tn*phoA* | T4(CP9.45) x CP9 |
| CP9.82 | *iroN*$_{ec}$:Tn*phoA*, active Tn*phoA* fusion | Kan$^R$exconjugant of CP9/pRT291 x MM294/pPH1J1 |
| CP82 | *iroN*$_{ec}$:Tn*phoA* | T4(CP9.82) x CP9 |

| Plasmid | | |
|---|---|---|
| p45.1 | 21kb*Sal1/SacI* fragment from CP45 containing the leftward 5.0kb of Tn*phoA* (active fusion) and a portion of *iroN*$_{ec}$ cloned into pBS II SK (-) | |
| p45.3 | 22kb *Cla1/Xba1* fragment from CP45 containing the rightward 6.7kb of Tn*phoA* (non-fusion) and a portion of *iroN*$_{ec}$ cloned into pBS II SK (-) | |
| p82.1 | 22kb *BamH1/Sac1* fragment from CP82 containing the leftward 5.0kb of Tn*phoA* (active fusion) and a portion of *iron*$_{ec}$ cloned into pBS II SK (-) | |

To establish that these mutants were isogenic derivatives of CP9, a series of studies were performed to exclude the possibility that cryptic mutations were acquired during the mutagenesis procedure.

Transduction of the transposon insertion back into the wild-type strain CP9 was accomplished using the bacteriophage T4 resulting in transductants CP45 and CP82. Whole cell DNA preparation, restriction enzyme (New England Biolabs, Beverly, MA.) mediated DNA digestion and Southern hybridization using PCR generated radioactive probes was performed as described (Russo et

al., 1993, supra). Primers 63 (SEQ ID NO:2) and 64 (SEQ ID NO:3) were used to amplify a 4.0kb internal fragment of Tn*phoA* (contained in pRT291), which were used to probe for Tn*phoA* insertions. Southern analysis of *Bgl II* digested whole cell DNA containing Tn*phoA* produces a 2.8kb internal fragment and two variable junctions fragment per copy. Lipopolysaccharide (LPS), capsular polysaccharide, and outer membrane protein profiles were determined as described (Russo et al., 1995, *Infec. Immun.*, 63:1263-1269).

In the mutants CP9.45 and CP9.82, and their respective T4 generated transductants, CP45 and CP82, it was established that the mutants had a single transposon insertion, that transduction of this insertion back into the wild type strain (CP9) resulted in a derivative that possessed the same degree of increased expression in urine relative to L-B medium as the original mutant, that in the transductants the transposon was physically in the same location as in the original mutant and that alterations in capsule, lipopolysaccharides and outer membrane profiles did not occur. This data establishes that CP45 and CP82 are isogenic derivatives of CP9 and that the gene into which TnphoA was inserted ($iroN_{ec}$) has increased expression in urine.

Subclones of the gene loci 5' to the Tn*phoA* insertions in CP9.45($iroN_{ec}$) were obtained by restricting whole cell DNA with *BamHI* or *SalI*, which recognizes a site located 3' to the kanamycin resistance gene in Tn*phoA* plus *SacI* (CP45, CP82) which does not have restriction sites within Tn*phoA*. Ligation of this restriction into pBSII SK(_), electroporation into XL1 Blue (Stratagene, Lajolla, CA) and selection of ampicillin (200$\mu$g/ml) and kanamycin (40$\mu$g/ml) resistant transformants resulted in the identification of the

subclones p45.1 and p82.1 (Table 1). A subclone of the *iroN*$_{ec}$ gene locus 3' to the Tn*phoA* insertion in CP9.45 was obtained by restricting whole cell DNA with *ClaI* which recognizes a site 5' to the kanamycin resistance gene in Tn*phoA* and *XbaI* which does not possess a restriction site within Tn*phoA*. Ligation of this restriction into pBSII SK(-), electroporation into XL1 Blue (Stratagene, Lajolla, CA), and selection of ampicillin and kanamycin resistant transformants resulted in the identification of the subclone p45.3.

DNA sequence was determined by the dideoxy chain termination method of Sanger (Sanger, et al, 1977, *Proc. Natl. Acad. Sci.* 85:5463-5467) using the gene subclones p45.1, p45.3 and p82.1 as the DNA templates. DNA sequencing of the gene subclones p45.1 and p82.1 initially utilized a Tn*phoA* fusion joint primer (SEQ ID NO:4) which established the location for a given Tn*phoA* insertion. Sequencing of the gene subclone p45.3 initially utilized the Tn*phoA* primer (SEQ ID NO:5). Subsequent DNA sequence was determined using primers derived from the deduced sequences of the gene subclones. A consensus sequence for *iroN*$_{ec}$ was generated by assembling and editing the DNA sequence obtained from 34 overlapping but independent sequencing reactions using AssemblyLIGN 1.0.2 (Oxford Molecular Group, Beaverton, OR). Both strands of the gene were sequenced. The sequence of *iroN*$_{ec}$ is disclosed in SEQ ID NO:1 (accession no. AF135597). Sequence analysis, comparisons, and CLUSTAL alignments were performed, in part, using MacVector (version 6.0, Oxford Molecular Group, Beaverton, OR). Comparisons were also performed via BLAST analysis of the non-redundant GenBank+EMBL+DDBJ+PDB sequences. SignalP V1.1 was used for identification of signal sequences (Perez-Trallero

et al., 1993, *Eur. J. Clin. Microbiol. Infect. Dis.*, 12:349-351).

The DNA sequence indicates that this is a novel gene for *E. coli*. The predicted protein ($IroN_{ec}$) consists of 725 amino-acids, a putative molecular weight of 79,380 kDa, an estimated pI of 5.68, and the first 24 a.a. represent a putative signal sequence. Comparison with entries in GenBank revealed that it was most homologous with $iroN_{se}$ (77% nucleotide homology), a recently identified catecholate siderophore receptor in *Salmonella enterica*. (Baumler, et al, 1998, 180:1446-1453). In *S. enterica*, $iroN_{se}$ is part of a five gene cluster that includes *iroBCDEN*. The specific functions of the gene products for *iroBCDE* remains unclear and a gene encoding the cognate siderophore for $IroN_{se}$ has yet to be identified. However, a putative Fur DNA binding site is present 5' to both $iroN_{se}$ (86 bases) and $iroN_{ec}$ (100 bases) with 15/19 bases conserved. Further, 184 bases 5' to the start site of $iroN_{ec}$ are 137 bases which were 87% (120/137) homologous to the *Salmonella* insertion sequence IS*1230* (bases 1-137, accession no. AJ000635). These bases correspond to the first 137 bases from this IS3-like element, the first 39 of which consist of an imperfect inverted repeat. (Collighan, et al, 1997, *FEMS Micobiol. Lett.* 154:207-213). This IS element was not present 5' to $iroN_{se}$. Taken together, this data suggests that $iroN_{ec}$ may have been acquired from *S. enterica* via IS-element mediated horizontal transfer. $iroN_{ec}$ possessed a lesser degree of homology with the catecholate siderophore receptors from *Pseudomonas aeruginosa* (*pfeA*), *Bordetella pertussis* (*bfeA*), and the *fepA* siderophore receptor from *E. coli*.

The deduced protein sequence identity and similarity of $IroN_{ec}$ determined by pairwise amino acid alignment using the program CLUSTAL to $IroN_{se}$, FepA, PfeA, and BfeA is summarized in Table 2. Further, sequence analysis did not reveal any premature stop codons.

Table 2

Deduced protein sequence identity and similarity of $IroN_{ec}$ to $IroN_{se}$ FepA, and BfeA

| | Homologue | % Identity | % Similarity |
|---|---|---|---|
| *Salmonella enterica* | $IroN_{ec}$ | 82% | 91% |
| *Escherichia coli* | FepA | 52% | 69% |
| *Pseudomonas aeruginosa* | PfeA | 52% | 69% |
| *Bordetella pertussis* | BfeA | 53% | 68% |

Additional sequence analysis of DNA 5' to *iroN*$_{ec}$ has demonstrated that this gene is 1.6kb 3' to the *prs* gene cluster, which encodes the class III PapG adhesin (Figure 1). Further, although the precise genomic organization of the region 5' to the *prs* operon has not been determined, the molecular usher for the F1C fimbria ((Klemm, et al, 1995, *J. Bacteriol.* 177:621-627) *focD*, has also been identified approximately 15.6kb 5' to the deduced start site for *iroN*$_{ec}$ (Figure 1). From left to right in figure 1 are shown: 11.8 kb sequence represented by the dotted line consists of the 5' portion of the *iroD*$_{ec}$ gene and genome 5' to it; a portion of the *iroD*$_{ec}$ gene (0.6kb), *iroE*$_{ec}$ (0.9kb) and

*$iroN_{ec}$*(2.2kb); a putative Fur binding site which is 100 bases 5' to the deduced start site of $iroN_{ec}$; a portion of an IS12230-like element (bp 1-137, accession number AJ000635) which is 184 bp 5' to the deduced start site of *$iroN_{ec}$*; and a 14 kb sequence represented by the dotted line (not to scale relative to solid line). The most distal boundary of this sequence consists of the 3' portion of focD, the molecular usher for the FIC fimbria. The arrow below the solid line indicate the direction of transcription.

**Example 2**

This embodiment describes the phylogenetic distribution of *$iroN_{ec}$* amongst various isolates of *E. coli*. A 667 base pair internal DNA probe (AF135597 bases 1729-2396) was generated from *$iroN_{ec}$* which did not share any homology with *fepA*. This probe was used in a dot-blot assay as described below to detect for the presence of homologous *$iroN_{ec}$* sequence.

DNA was prepared from relevant strains by boiling cells from overnight growth in L-B medium (1ml concentrated to 200μl of sterile $H_2O$) at 105°C for 10 minutes. The supernatant was saved and used for analysis. Nytran membranes (Schleicher & Scheull, Keene, New Hampshire) were pre-wet in 6X SSC for 10 minutes and dotted with 3 μl of denatured DNA preparation from each strain in triplicate. The membrane was subsequently placed on filter paper saturated with denaturing solution (0.4N NaOH, 0.6M NaCl) followed by neutralizing solution (1.5M NaCl, 0.5M Tris HCl, pH 7) for 1 minute each, then air-dried, and UV cross-linked with 1200 Joules (UV Stratalinker 2400, Statagene, La Jolla, CA). Primers 192 (SEQ ID NO:6) and 197 (SEQ ID NO:7) were used to amplify a 0.67kb internal fragment of *$iroN_{ec}$*

(contained in p45.3) that did not share any homology with *fepA*. Aqueous hybridization was performed under high stringency conditions (65°C). Results were scored as either zero (no hybridization), $1^+$, or $2^+$. Under these conditions the negative control strains (HB101, XL-1 Blue) were consistently scored as zero and the positive control strain CP9 as $2^+$. Experimental strains consisted of; Group 1: fourteen unique fecal isolates that had been previously established not to contain *pap*, *hly*, or *cnf-1* (Johnson, et al, 1998, *J. Infect. Dis.* 177:1120-1124), Group 2: five unique fecal isolates that possessed some combination of *pap*, *hly*, or *cnf-1*, Group 3: twenty unique first-time UTI isolates (Russo, et al, 1995, *J. Infect. Dis.* 172:440-445), Group 4: fifteen unique recurrent UTI isolates (Russo, et al, 1995, *J. Infect. Dis.* 172:440-445), Group 5: twenty-one blood isolates obtained from patients hospitalized at Erie County Medical Center (Buffalo, NY), Group 6: all 35 UTI isolates, and Group 7: all 56 clinical isolates. Group 1 was most representative of non-pathogenic or commensal strains and therefore was used in statistical comparisons against the clinical isolate Groups 5, 6, and 7. Fisher's exact test was used for the comparison of fecal versus clinical isolates for the presence of $iroN_{ec}$ DNA sequence via dot-blot assay.

The results from these studies are summarized in Table 3.

Table 3

Phylogenetic distribution of *ironN$_{ec}$* homologous DNA sequence

| Dot-Blot Score | Group 1 | Group 2 | Group 3 | Group 4 | Group 5 | Group 6 | Group 7 |
|---|---|---|---|---|---|---|---|
| 0 | 43% | 20% | 5% | 7% | 10% | 6% | 7% |
| $1^+$ | 43% | 0% | 50% | 27% | 57% | 40% | 46% |
| $2^+$ | 14% | 80% | 45% | 67% | 33% | 54% | 46% |
| any- | 57% | 80% | 95% | 93% | 90% | 94% | 93% |

Dot-Blot Score: 0= no homology, $1^+$ = positive homology, $2^+$ = maximal homology, any$^+$ = the combination of $1^+$ and $2^+$

Group 1: fourteen unique fecal isolates that had been previously established not to contain *pap*, *hly*, or *cnf-1* and therefore are most representative of non-pathogenic strains.

Group 2: five unique fecal isolates that possessed some combination of *pap*, *hly*, or *cnf-1* and therefore most likely represent pathogenic strains

Group 3: twenty unique first-time UTI isolates

Group 4: fifteen unique recurrent UTI isolates

Group 5: twenty-one blood isolates obtained from patients hospitalized at Erie County Medical Center (Buffalo, NY).

Group 6: All 35 UTI isolates (Groups 3 and 4).

Group 7: All 56 clinical isolates (Groups 3, 4 and 5).

[a]Fisher's exact test was used for proportions. All comparisons are versus Group 1.

[b]P=0.039, [c]P=0.004, [d]P=0.003, [c]P=>0.10(NS), [f]P=0.01, [g]P=0.03

In summary, forty three percent of 14 fecal isolates (Group1, negative for *pap*, *hly*, or *cnf1* and

therefore most representative of non-pathogenic strains) did not possess DNA sequence homologous to *iroN* $_{ec}$. In contrast, only 20% of 5 fecal isolates (Group 2, positive for some combination of *pap*, *hly*, or *cnf1* and therefore most likely pathogenic strains), 5% of 20 first time UTI isolates (Group 3), 7% of 15 recurrent UTI isolates (Group 4), and 10% of 21 blood isolates (Group 5) were negative for *iroN*$_{ec}$ homologous sequence under high stringency conditions. The differences between Group 1 versus either Group 5, Group 6 (all UTI strains, Groups 3,4), or Group 7 (all clinical isolates, Groups 3,4,5) were statistically significant (P = 0.039, P = 0.004, P = 0.003 respectively). In summary, this data demonstrates that DNA sequence homologous to *iroN*$_{ec}$ is significantly less prevalent in fecal isolates without the virulence genes *pap*, *hly*, or *cnf1* than clinical isolates.

**Example 3**

This embodiment describes the growth of CP9 and CP82 and the expression of *iroN*$_{ec}$ in human urine, human ascites and blood.

Ex vivo growth in human urine

Human urine from subjects who had and who never had experienced a UTI was used for studies assessing growth of strains *ex vivo*. The strain to be tested was grown overnight in 2 ml of L-B medium ± kanamycin 40μg/ml. The next day, the bacterial cells were diluted into urine to achieve a starting concentration of approximately 1.0 x $10^{2-3}$ cfu/ml, since this titer is at the lower end of the spectrum for what is considered significant for UTI in symptomatic young women (Stamm, et al, 1982, *N. Engl. J. Med.* 307:463-468). For $Å_{600}$ growth curves, a starting $Å_{600}$ of about 0.03 was used. During incubation at 37°C,

aliquots were removed at intervals and either the $\mathring{A}_{600}$ was determined or the bacterial titers were established by plating 10-fold serial dilutions in 1X phosphate-buffered saline in duplicate on appropriate media.

Expression of $iroN_{ec}$ in urine

Expression of *iroN*$_{ec}$ was increased in human urine relative to L-B. Since the composition of human urine has the potential to be variable, assays were performed using 17-29 independent urines collected from 10 different individuals. Five of these individuals were women with a prior history of UTIs. For CP9.82, the mean fold and median fold increase in expression of $IroN_{ec}$ was 27.2±5.0 and 19.0 respectively, the range being 2.4-132. Although there was variance in the degree of increased expression from urine to urine, increased expression was seen in all urines evaluated. The degree of expression of *iroN*$_{ec}$ was similar in urines from individuals with and without a prior history of UTI. The 17-29 independent urines used were filter sterilized and stored at 4°C prior to use. To determine if the processing of urine affected gene expression, assays were performed in parallel using urines that were either 1) fresh and unfiltered, 2) fresh and filtered with a .22 micron filter, or 3) filtered and stored at 4°C. The expression of *iroN*$_{ec}$ was similar, regardless of how the urine was processed (data not shown).

Growth of CP9 (w.t.) and CP82 (*iroN*$_{ec}$) was evaluated in multiple independent urines via both enumeration of bacterial titers and $\mathring{A}_{600}$. Growth of CP82 (5 urines) was equivalent to their wild-type parent CP9 (data not shown).

Expression of *iroN*$_{ec}$ in human ascites and blood

The expression of various virulence traits may vary depending on the site of infection. Therefore *iroN*$_{ec}$, expression was evaluated in human blood and ascites, two

additional body fluids which extraintestinal *E. coli* isolates commonly infect. To assess gene expression in human ascites and blood, filter sterilized ascites (peritoneal fluid) was obtained from a patient hospitalized at Erie County Medical Center, divided into multiple aliquots, and frozen at -80°C. Blood was used fresh and was obtained from a single donor. It was collected in sterile, 8.3ml vacutainer tubes which contained 1.7ml of sodium polyanetholesulfonate (0.35%) and NaCl (0.85%) (non-bactericidal) as the anti-coagulant. For blood assays the bacterial cells were washed x 2 at 4°C with 4ml of 0.1M Tris (pH 9.8), 0.001M $MgCl_2$ buffer and resuspended in a total volume of 2ml. For ascites assays the bacterial cells were concentrated via centrifugation and the resultant pellet was resuspended in 2ml of 0.1M Tris (pH 9.8), 0.001M $MgCl_2$ buffer. Aliquots were removed x 2 and cfu/ml were determined via serial 10-fold dilutions. Bacterial cells were subsequently permeabilized by adding 100$\mu$l of 0.1% SDS and 200$\mu$l of chloroform, vortexed x 10 seconds and kept on ice. A fluorescent assay was performed because red blood cells ± hemoglobin present in blood could not be reliably separated from bacterial cells. Their presence interfered with the colorimetric assays described above for measuring alkaline phosphatase activity. Assays were performed in a 48-well tissue culture plate. Each assay mixture consisted of 1ml of Tris buffer, 50$\mu$l of bacterial cell extract, and 50$\mu$l (0.01M) of the fluorescent substrate (4-methylumbelliferonephosphate). Samples were read using a fluorescence multi-well plate reader (CytoFluor II, PerSeptive Biosystems, Framingham, MA) at an excitation setting of 360 nm, an emission setting of 460 nm, and a gain of 80 for 15 cycles. The net sample rate in blood or ascites relative to that in L-B broth established the fold induction. The net sample rate / ml (SR) =

{((fluorescence cycle B - fluorescence cycle A over the linear portion of the curve)/(elapsed time)) x 20} - (CP9 SR). Specific activities were determined by dividing net sample rates by cfu/ml. The sensitivity of the colorimetric and fluorescent assays was established to be similar.

The expression of *iroN*$_{ec}$ in blood and ascites was as follows. The mean fold increase in blood was 65.8±6.7 while the mean fold increase in ascites was 207±27. Although there was some variance in expression compared to urine, it should be noted that ascites and blood were obtained from single individuals.

**EXAMPLE 4**

This embodiment describes the regulation of the expression of *iroN*$_{ec}$ under various environmental conditions.

For osmoregulation studies, modified Davis medium was used with variable concentrations of either NaCl (0.05M-0.7M) or urea (0.05M-0.7M). Some gene regulation studies utilized urine to which exogenous Fe (0.1mM) or glucose (0.5%) was added. M9 minimal medium was also utilized in gene regulation studies. Fe was chelated from M9 medium by mixing 200ml of medium with 21.2 grams of washed (with 1L $dH_2O$ x 2) iminodiacetic acid (Chelex 100, Sigma, St. Louis, MO.) for 90 minutes followed by filter sterilization. Siderophore production was determined using the Arnow assay as described (Schwyn, et al, 1987, *Anal. Biochem.* 160:47-56) and was concomitantly measured to confirm that the Fe concentration was limiting when chelated. As expected, siderophore production increased from 3.1μM/$\mathring{A}_{600}$ in the presence of Fe (0.1mM) to 10.7μM/$\mathring{A}_{600}$ when Fe was chelated. The effect of pH on expression was determined using pooled urine whose pH was adjusted with either HCl

or NaOH to achieve pHs of 5.0, 6.0, and 7.0. For a given experiment, assays were performed in triplicate and experiments were repeated at least once. Results were presented as the ratio of reporter gene expression in urine relative to L-B medium. For all of these studies quantitative alkaline phosphatase assays were performed as described above.

Iron: Sequence analysis of *iroN*$_{ec}$ strongly supports that this gene codes for a catecholate siderophore receptor, and therefore the role of Fe in the regulation of *iroN*$_{ec}$ was evaluated. Expression of *iroN*$_{ec}$ was measured when CP82 (*iroN*$_{ec}$) was grown in M9 minimal media in which Fe was either chelated or added exogenously. Compared to M9 medium plus Fe, *iroN*$_{ec}$ expression was increased 20.8-fold when CP82 was grown in Fe chelated M9 medium. Further, the addition of exogenous Fe to 3 independent human urines suppressed the increased expression of *iroN*$_{ec}$ relative to L-B medium (mean 64-fold decrease in *phoA* activity). These experimental findings, in combination with the identification of a Fur binding sequence 5' to the start of *iroN*$_{ec}$, suggest that *iroN*$_{ec}$ is Fur- regulated and that Fe is limited in urine.

pH: Although the pH of normal urine most commonly ranges from 5.5 - 6.5, values from 5.0 to 8.0 can occur. Therefore the effect of pH 5.0, 6.0, and 7.0 on the expression of *iroN*$_{ec}$ in human urine was evaluated. The expression of *iroN*$_{ec}$ was completely suppressed at pH 5.0 but unaffected at pH 6.0 and 7.0 with induction ratios of 0.23, 32, and 34 measured respectively. Therefore, urinary pH can affect gene expression of *iroN*$_{ec}$.

Thus, low Fe concentrations increase the expression of *iroN*$_{ec}$, and its expression is suppressed at a urinary pH 5.0, but unaffected by limiting concentrations of amino acids, nucleotides, or glucose.

**EXAMPLE 5**

This embodiment illustrates the use of the $IroN_{ec}$ for antigenic purposes. The entire protein encoded by *iroN*$_{ec}$ or antigenic fragments thereof can be used in vaccine formulations. Surface exposed epitopes of $IroN_{ec}$ can be identified by methods well known to those skilled in the art. For vaccine development, the $IroN_{ec}$ protein may be purified from the bacteria or may be purified from host containing a recombinant vector which expresses *iroN*$_{ec}$. The antigenic formulation may be introduced into the human or animal to be vaccinated by standard techniques well known to those skilled in the art.

To illustrate this embodiment, the $IroN_{ec}$ was purified following the cloning of the gene into an expression vector. The purified protein was used to elicit antibodies in mice. Immunized mice were challenged by the homologous strain of *E.coli* and protective abilities of this protocol determined.

<u>Cloning of the *iroN*$_{ec}$</u>

For PCR-mediated amplification, the following 2 primers were designed for the entire *iroN*$_{ec}$ gene, excepting its signal sequence (2083 base pairs). These primers were derived from the *iroN* sequence (SEQ ID NO:1). The forward primer CGCGCGCGGATCCGACGAGACTCTGGTGGTGGA (SEQ ID NO:9) and the return primer CGCGCGCAAGCTTGAATGATGCGGTAACTCCGG (SEQ ID NO:10) were used. A single band of the expected size was PCR amplified from CP9 chromosomal DNA. The DNA was cleaned and ligated into the Kanamycin resistant pet28a T7/his-tag expression vector. The pet28a::*iroN* construct was electroporated into XL1 Blue cells and selected for on LB plates containing Kanamycin. The *iroN*$_{ec}$ gene in the selected clone was confirmed to be

correct by DNA sequencing. This excluded the possibility that an error was introduced into the cloned *iroN*$_{ec}$ during PCR amplification. The clone was subsequently electroporated into the expression cell line AD494 (DE3) pLysS for over-expression of IroN$_{ec}$.

AD494 (DE3) pLysS::pet28a::*iroN*$_{ec}$ was grown overnight in LB media plus Kanamycin. The next morning, 1 ml of the overnight culture was transferred into 11ml LB media plus Kanamycin and grown at 37°C for 2.5 hours, shaking. IPTG was added to a final concentration of 1mM to induce the expression of IroN$_{ec}$. One ml aliquots of the induced culture as well as an uninduced control culture were taken in 30 minute intervals. The samples were prepared for gel electrophoresis and run on a 7.5% SDS-PAGE gel (Figure 2A). This figure shows the increased expression of IroN$_{ec}$ in the induced culture, migrating at approximately 80 kDa . This size is close to the deduced size of IroN$_{ec}$ (79.4 kDa) based on its DNA sequence. IroN$_{ec}$ was subsequently purified using TALON cobalt-based Immobilized Metal Affinity Chromatography and eluted under denaturing conditions using a 6M urea elution buffer at a pH between 5.1 and 5.3. A Western blot using the T7-Tag antibody specific to the recombinant protein was done. An intact IroN$_{ec}$ is represented by the primary band (Figure 2B). The minor bands below IroN$_{ec}$ represent break-down products of IroN$_{ec}$ since these bands are all recognized by the T7-tag antibody which is specific to recombinant IroN$_{ec}$. This established that IroN$_{ec}$ was successfully purified. Purified IroN$_{ec}$ was stored at -20°C until further use.

The successful expression of IroN$_{ec}$ confirmed that *iroN*$_{ec}$ encoded as a full length product. Further, the over-expressed and purified IroN$_{ec}$ was used as an immunogen in mice to determine if it was antigenic and elicited an antibody response.

Purified $IroN_{ec}$ elicits a strong antibody response in mice

A putative vaccine candidate must be immunogenic. To establish that this is the case for $IroN_{ec}$, mice were immunized with purified $IroN_{ec}$ and sera pre and post-immunization were obtained and tested as follows.

**a) Immunization protocol:**

Balb C mice were used in an immunization protocol that was performed over a 5 week period. Mice were divided into 3 groups: Group 1, controls immunized with buffer only (N=15); Group 2, animals immunized with a total of 70$\mu$g of $IroN_{ec}$ (N=12); and Group 3, animals immunized with a total of 150$\mu$g of $IroN_{ec}$ (N=15). Purified $IroN_{ec}$ was injected subcutaneously on days 1, 15 and 30 in a total of 200$\mu$l for the first 2 immunizations and in 100$\mu$l for the last immunization. No adjuvant was used. Sera were collected pre-immunization on day 1 and subsequently on days 22 and 36. The pre and post-immunization sera from these animals were evaluated by enzyme linked immunosorbent assay (ELISA) assay to assess for the development of antibodies directed against $IroN_{ec}$.

**b) ELISA assay for detection of antibodies directed against IroNec:**

ELISA assays were performed using Immulon 2 HB plates (DYNEX), coated with 75ng of purified IroNec/well. The serum dilution was 1/1000. The conjugate used was Peroxidase-Labeled Goat anti Mouse IgG + IgM at a concentration of 1/10,000. $IroN_{ec}$ was adsorbed to the plate overnight. In the morning, the plate was washed and blocked using PIERCE Superbloc. After blocking, the plate was washed again and incubated with the diluted sera for 2 hours. The plate was washed

again and incubated with the diluted conjugate for 1 hr before it was developed. Readings were measured at 450nm on an automated ELISA plate reader.

The results of this assay are summarized in Table 4. As can be seen, a significant increase in antibodies directed against $IroN_{ec}$ developed in mice immunized with $IroN_{ec}$ (Groups 2 and 3) when compared with non-immunized controls (Group 1). Further, every animal in Group 2 and 3 had an antibody response of a similar magnitude. However, there was no difference in the magnitude of antibody response between the animals immunized with a total of 70$\mu$g of $IroN_{ec}$ (Group 2) when compared to animals immunized with a total of 150 $\mu$g of $IroN_{ec}$ (Group 3). These results demonstrate that $IroN_{ec}$ is antigenic, a critical property of a vaccine candidate.

Table 4:
Immunization with $IroN_{ec}$ elicits a strong antibody response in mice

| Immunizing regimen | ELISA $OD_{600}$ mean±SEM | Post/Pre ratio |
|---|---|---|
| Controls | 0.043±0.0007 (pre)<br>0.090±0.010 (post) | 2.1 |
| Immunized with 70$\mu$g of $IroN_{ec}$ | 0.084±0.015 (pre)<br>1.31±0.022 (post) | 15.6 |
| Immunized with 150$\mu$g of $IroN_{ec}$ | 0.099±0.008 (pre)<br>1.38±0.015 (post) | 14.7 |

**Example 6**

This embodiment demonstrates that immunization with $IroN_{ec}$ protects mice against challenge with the *Escherichia coli* strain CP9. Having established that immunization with $IroN_{ec}$ results in antibody production

directed against $IroN_{ec}$, experiments were performed to evaluate whether this antibody response was protective against challenge with the homologous *E. coli* strain CP9. For initial studies an intra-peritoneal challenge model was used.

**a) Intraperitoneal challenge model:**

Both male and female mice (18-22 grams) underwent the immunizing regimen described above using purified $IroN_{ec}$ as the immunogen. The non-immunized control Group 1 consisted of 15 animals. Group 2 (N=12) and Group 3 (N=15) were immunized with a total of 70$\mu$g and 150$\mu$g of purified $IroN_{ec}$ respectively. Sera were obtained prior to and after immunization. Twelve days after the third and final immunizing dose of $IroN_{ec}$ was administered, animals underwent intra-peritoneal challenge with four different titers of the *E. coli* strain CP9. This infection model results in a systemic infection that may be lethal, depending on the magnitude of the challenge titer. For these experiments, challenge titers were utilized that would result in a 50-75% mortality rate in control animals. In this manner, a protective effect of immunization with $IroN_{ec}$ could be identified. The measured endpoints of this study were 1) death and 2) hepatic and splenic bacterial titers. Animals were observed post- bacterial challenge. Upon death, the liver and spleen were immediately removed, homogenized, and titered for bacterial counts via serial 10-fold dilutions. If an animal was still alive 18 hours after bacterial challenge it was sacrificed and bacterial titers of the liver and spleen were performed as described above. Since the antibody response against $IroN_{ec}$ was the same for Groups 2 and 3, these groups were pooled for the analysis presented below.

When antibody response was evaluated by ELISA, none

of the control animals possessed significant pre-existing antibody to $IroN_{ec}$, nor did a significant antibody response occur after sham immunization. In contrast, all of the animals immunized with $IroN_{ec}$ developed a significant antibody response against it. The results are presented in Table 5.

Table 5

Immunization with $IroN_{ec}$ decreases mortality in mice challenged intraperitoneally with live *Escherichia coli* strain CP9

| Immunizing regimen | no. dead/no. Injected (%) | $LD_{50}$[&] |
|---|---|---|
| Negative controls not immunized with $IroN_{ec}$ | | |
| $2.5 \times 10^6$ cfu[#] | 1/3 (33) | $3.83 \times 10^6$ cfu |
| $8.1 \times 10^6$ | 3/4 (75) | |
| $2.5 \times 10^7$ | 4/4 (100) | |
| $8.1 \times 10^7$ | 4/4 (100) | |
| Immunized animals (immunized with $IroN_{ec}$) | | |
| $2.5 \times 10^6$ | 1/6 (16.6) | $7.84 \times 10^6$ cfu |
| $8.1 \times 10^6$ | 3/6 (50) | |
| $2.5 \times 10^7$ | 5/7 (71.4) | |
| $8.1 \times 10^7$ | 7/7 (100) | |

& = bacterial challenge dose needed for death in 50% of mice

# = bacterial colony forming units

As shown in Table 5, immunization with $IroN_{ec}$ both resulted in a decreased mortality in immunized animals (higher $LD_{50}$ dose, $7.84 \times 10^6$ cfu ) when compared to unimmunized controls (lower $LD_{50}$ dose, $3.83 \times 10^5$ cfu). Further, immunized control animals had diminished growth of CP9 in the liver and spleen when compared to unimmunized control animals (Table 6).

Table 6

<u>Immunization with $IroN_{ec}$ diminishes bacterial growth in the liver and spleen of mice challenged intraperitoneally with live *Escherichia coli* strain CP9</u>

| Immunizing regimen | Hepatic & Splenic growth/no. Injected (%) | $GD_{50}$[&] |
|---|---|---|
| Negative controls (not immunized with $IroN_{ec}$) | | |
| $2.5 \times 10^5$ cfu[a] | 1/3 (33) | |
| $8.1 \times 10^5$ | 4/4 (75) | $2.58 \times 10^5$ cfu |
| $2.5 \times 10^7$ | 4/4 (100) | |
| $8.1 \times 10^7$ | 4/4 (100) | |
| Immunized animals (immunized with $IroN_{ec}$) | | |
| $2.5 \times 10^6$ | 1/6 (16.6) | |
| $8.1 \times 10^5$ | 3/6 (50) | $7.84 \times 10^5$ cfu |
| $2.5 \times 10^7$ | 5/7 (71.4) | |
| $8.1 \times 10^7$ | 7/7 (100) | |

& = bacterial challenge dose needed for growth of CP9 in the liver and spleen of 50% of the mice

# = bacterial colony forming units

In summary, immunization with purified $IroN_{ec}$ resulted in the development of antibodies directed against it. Further, these antibodies resulted in protection against subsequent bacterial challenge as shown by a diminished mortality rate and diminished growth of the model pathogen CP9 in liver and spleen.

From the foregoing, it will be obvious to those skilled in the art the various modifications in the above-described methods and compositions can be made without departing from the spirit and scope of the invention. Accordingly, the invention may be embodied in other specific forms without departing from the spirit or essential characteristics thereof. Present examples and embodiments, therefore, are to be considered in all respects as illustrative and not restrictive, and all changes which come within the meaning and range of equivalency of the specifications are therefore intended to be embraced therein.

What is claimed is:

1. An isolated polynucleotide comprising a polynucleotide sequence selected from the group consisting of:
(a) a polynucleotide having at least 90% homology to a polynucleotide encoding a polypeptide of SEQ ID NO:8 and
(b) a polynucleotide which is complementary to the polynucleotide of (a).

2. The isolated and purified polynucleotide of claim 1, wherein the polynucleotide in (a) has at least 95% homology to a polynucleotide encoding a polypeptide of SEQ ID NO:8.

3. The isolated and purified polynucleotide of claim 2, wherein the polynucleotide in (a) has at least 97% homology to a polynucleotide encoding a polypeptide of SEQ ID NO:8.

4. The isolated and purified polynucleotide of claim 3, wherein the polynucleotide in (a) is a polynucleotide encoding a polypeptide of SEQ ID NO:8.

5. The isolated and purified polynucleotide of claim 4, wherein the polynucleotide in (a) encoding a polypeptide of SEQ ID NO:8 has a sequence of SEQ ID NO:1.

6. A recombinant vector comprising the polynucleotide of claim 1.

7. A recombinant vector comprising the polynucleotide of claim 4.

8. A recombinant vector comprising a nucleotide sequence encoding one or more antigenic epitopes of $IroN_{ec}$.

9. A pure peptide of SEQ ID NO:8

10. A pure antigenic peptide, polypeptide or protein having one or more epitopes of $IroN_{ec}$, wherein $IroN_{ec}$ has the sequence of SEQ ID NO:8.

11. The peptide or protein of claim 9, wherein the peptide, polypeptide or protein is produced recombinantly.

12. An antigenic formulation comprising a pure peptide, polypeptide or protein having one or more epitopes of $IroN_{ec}$, wherein $IroN_{ec}$ has the sequence of SEQ ID NO:8.

13. The antigenic formulation of claim 13 further comprising a pharmaceutical carrier.

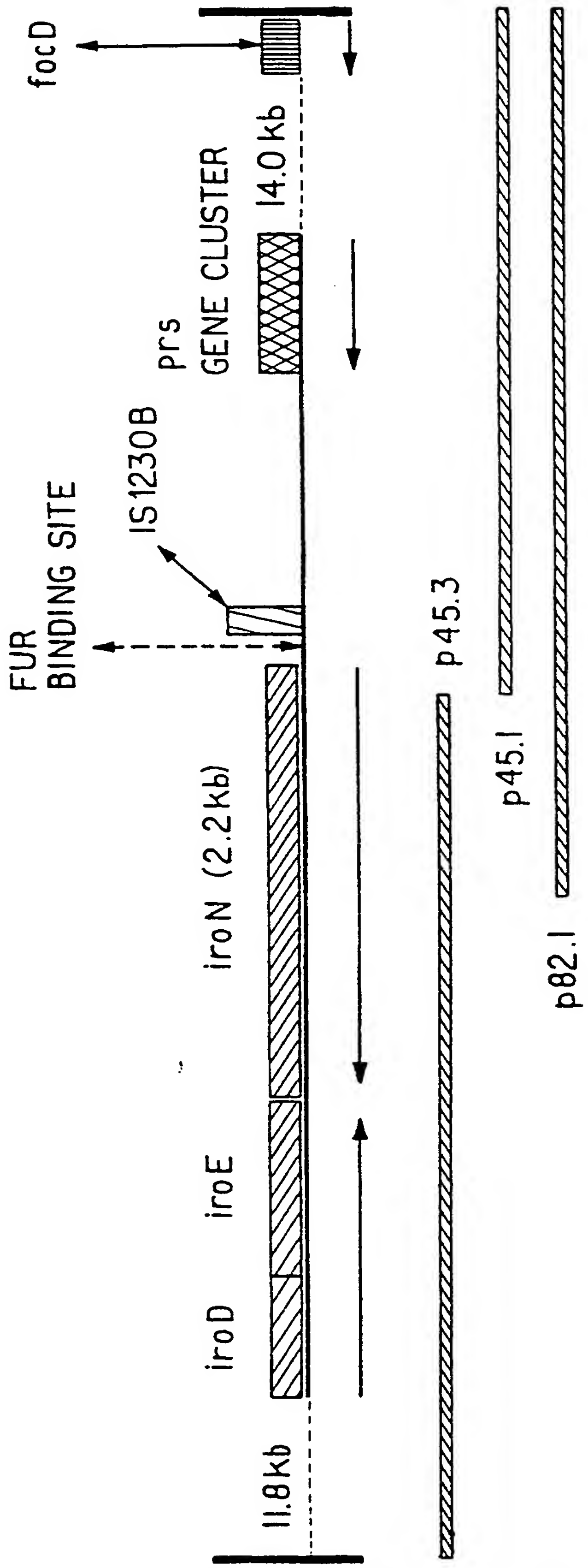
focD
14.0 kb
prs
GENE CLUSTER
IS1230B
FUR
BINDING SITE
iroN (2.2 kb)
iroE
iroD
11.8 kb
p45.3
p45.1
p82.1

FIG. 1

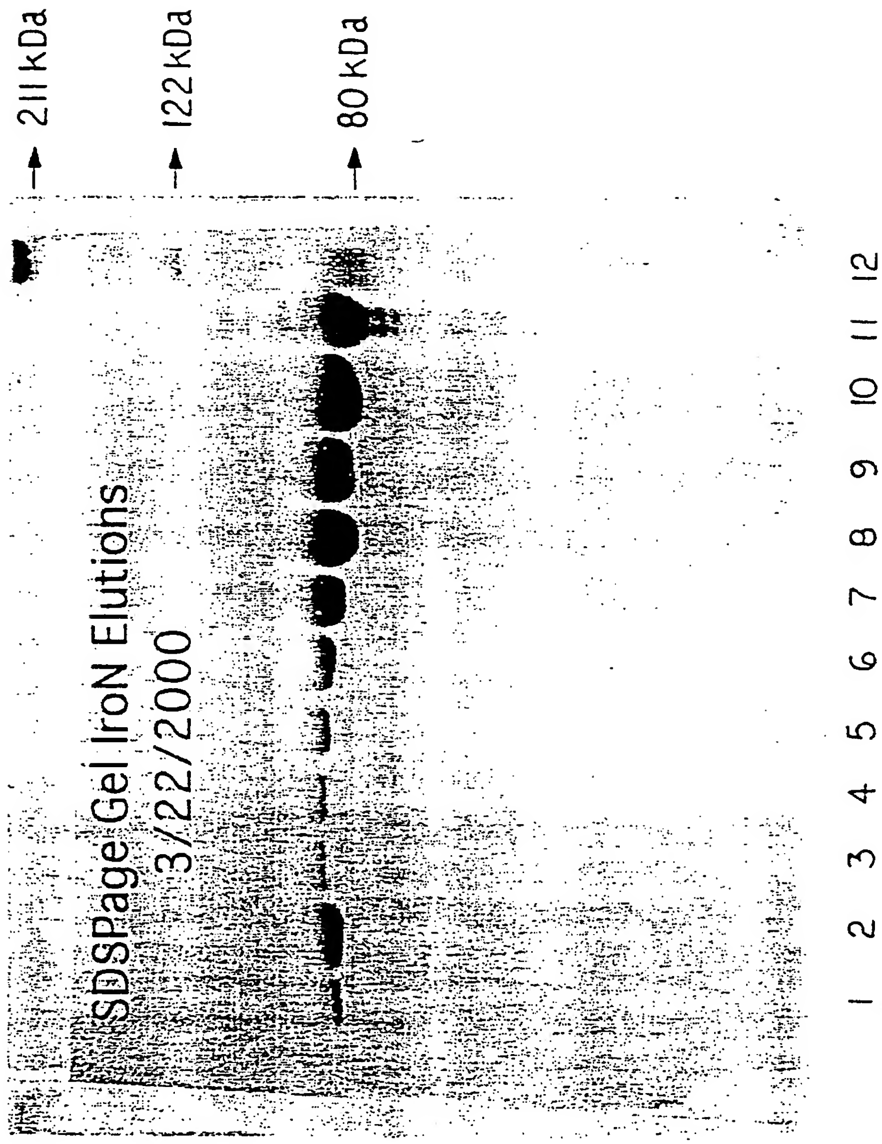
211 kDa
122 kDa
80 kDa
SDSPage Gel IroN Elutions
3/22/2000
1 2 3 4 5 6 7 8 9 10 11 12

FIG. 2A

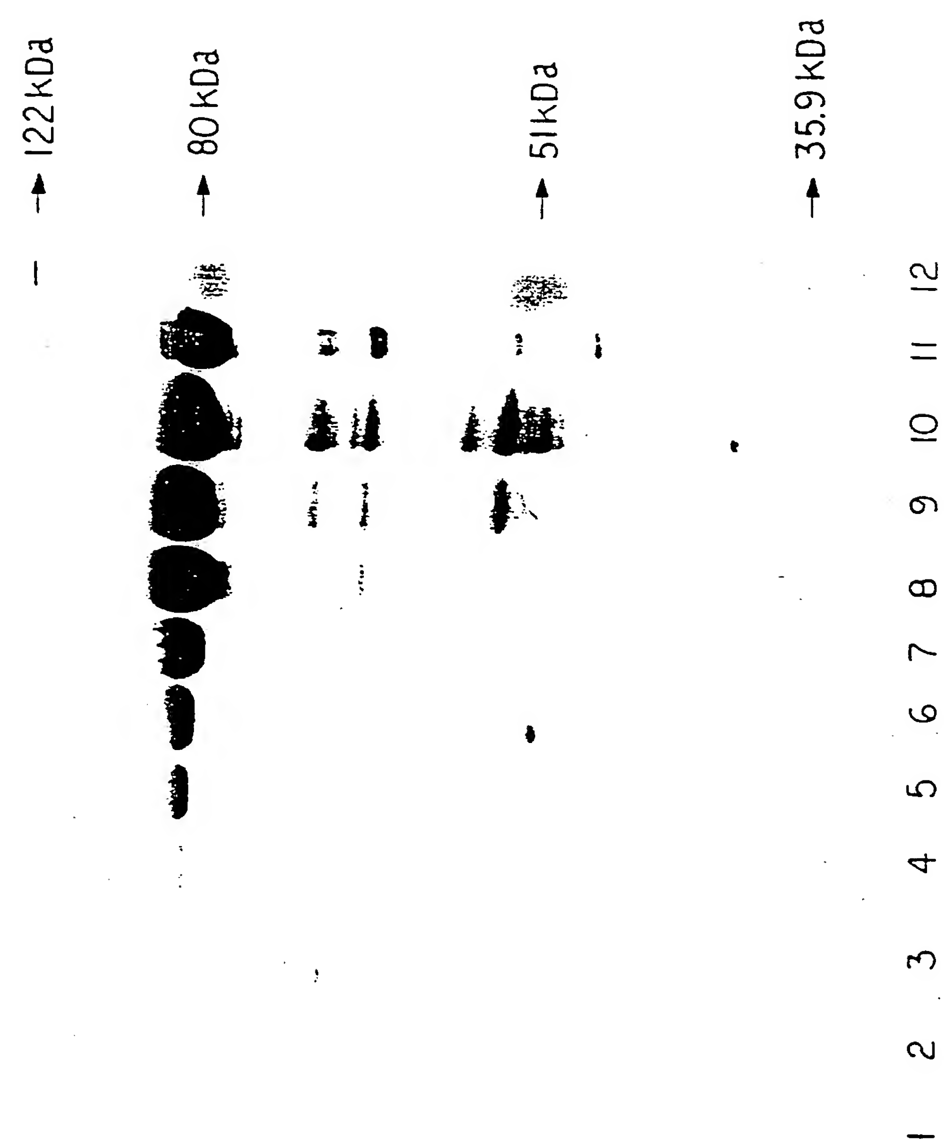


FIG. 2B

SEQUENCE LISTING

<110> Russo, Thomas A.
<120> Identification of A Vaccine Candidate from an Extraintestinal Strain of E. coli
<130> 11520.0214
<160> 5
<210> 1
<211> 2495
<212> DNA
<213> Escherichia coli
<220>
<400> 1

```
ctctgactcg cccgtcacga ctggaatcat gaaggcgttc aagttcagct   50
ttttgttgct cagtaatgaa tattttcatg ttggacagca tgatcttctg  100
ttgctgaaaa atcaagcatc ttcaatgatc acgggtatat atcaatttga  150
tatttatgaa aaatcaatag atttatcttg tgagggattg ctttttcttt  200
tgctaaatga tattaattat cattcttata taaaaacaat aatggcagta  250
agaaatatct ggcaaggatg tgagcttaac gatcaaacat tttaagttag  300
gcatttatta gggaatagga atgagaatta acaaaatcct ctggtcgcta  350
actgtgctcc tagttgggtt gaatagccag gtatcagtag ccaaatactc  400
cgacgatgat aatgacgaga ctctggtggt ggaagccacc gctgagcagg  450
tattaaaaca gcagccgggc gtgtcggtta ttaccagcga ggatattaaa  500
aagacccctc cggtaaacga cctttcagat attattcgta aaatgcctgg  550
tgttaatctt accggcaata gcgcctcggg cacacgcggt aataaccgcc  600
agatcgatat tcgtggtatg ggccggaaa acaccttaat tttaattgat  650
ggtgtaccgg tgacgtcacg taactccgtg cgttatagct ggcgtgggga  700
gcgtgatacc cgcggtgaca ccaactgggt gccaccggaa caggttgagc  750
gtattgaagt gatccgcggc cctgcggcgg cgcgctacgg ttcgggggcc  800
gccggggggg tggtgaacat cattaccaaa cgtcccacca acgactggca  850
cggttcgctg tcgttataca ccaaccagcc ggaaagtagc gaagagggcg  900
ctacgcgtcg cgccaatttc agccttagtg ggcctctggc tggtgatgct  950
cttaccacgc gtttgtatgg taacctgaat aaaacggatg ctgacagttg 1000
ggatattaat tctccggtcg gtacgaaaaa cgcagccggg catgaagggg 1050
tacgtaacaa agatattaac ggcgttgtct cgtggaaatt aaatccgcag 1100
cagattctcg atttcgaagt cggatatagc cgccagggga atatctatgc 1150
gggcgatacg cagaacagtt cttccagtgc agttaccgaa agcctggcaa 1200
aatccggcaa agagacgaac cgcctgtacc gacagaatta tggcattacg 1250
cataatggta tctgggactg ggacaaagt cgctttggtg tttattacga 1300
gaaaaccaat aatacccgca tgaatgaagg attatccggc ggtggtgaag 1350
gacgtatttt agcgggtgaa aagtttacga ccaatcgcct gagttcctgg 1400
cgaaccagcg gtgagcttaa tattcctttg aatgtgatgg ttgatcaaac 1450
gctgaccgtt ggtgcagagt ggaaccgcga taagctcgat gatccttcct 1500
ctaccagcct gacggtgaat gacagagata tcagcggtat ttctggctct 1550
gctgcggatc gcagcagtaa aaatcattct caaatcagtg cgctgtatat 1600
tgaagataac attgagccgg ttcctggcac gaatatcatt cccggcctgc 1650
gctttgatta tctcagcgac tccggcggga acttcagccc cagtctgaat 1700
ctttcgcagg aattgggcga ttatttcaaa gtcaaagcag gggttgcccg 1750
aacctttaaa gccccaaacc tgtatcaatc cagtgaaggc tatctgctct 1800
actcgaaagg caatggctgt ccaaaagata ttacatcagg cgggtgctac 1850
ctgatcggta ataaagatct cgatccggaa atcagcgtca ataaagaaat 1900
tggactggag ttcacctggg aagattacca cgcaagtgtg acctacttcc 1950
gcaatgatta ccagaataag atcgtggccg gggataacgt tatcggcaa 2000
accgcttcag gcgcatatat cctcaagtgg cagaatggcg ggaaagctct 2050
```

```
ggtggacggt atcgaagcca gtatgtcttt cccactggtg aaagagcgtc 2100
tgaactggaa taccaatgcc acatggatga tcacttcgga gcaaaaagac 2150
accggtaatc ctctgtcggt catcccgaaa tatactatca ataactcgct 2200

taactggacc atcacccagg cgttttctgc cagcttcaac tggacgttat 2250
atggcagaca aaaaccgcgt actcatgcgg aaacccgcag tgaagatact 2300
ggcggtctgt caggtaaaga gctgggcgct tattcactgg tggggacgaa 2350
cttcaattac gatattaata aaaatctgcg tcttaatgtc ggcgtcagta 2400
atatcctcaa taaacagatc ttccgatctt ctgaaggggc gaatacctat 2450
aacgagccag gccgggctta ttatgccgga gttaccgcat cattc      2495
```

<210> 2
<211> 18
<212> DNA
<213> artificial sequence
<220>
<223> completely synthesized primer
<400> 2

gatcaagaga caggatga 18

<210> 3
<211> 18
<212> DNA
<213> artificial sequence
<220>
<223> completely synthesized primer
<400> 3

tgatcctcgc cgtactgc 18

<210> 4
<211> 15
<212> DNA
<213> artificial sequence
<220>
<223> completely synthesized primer
<400> 4

aatatcgccc tgagc 15

<210> 5
<211> 18
<212> DNA
<213> artificial sequence
<220>
<223> completely synthesized primer
<400> 5

catgttagga ggtcacat 18

<210> 6
<211> 18
<212> DNA

<213> artificial sequence
<220>
<223> completely synthesized primer
<400> 6

gacgccgaca ttaagacg 18

<210> 7
<211> 19
<212> DNA
<213> artificial sequence
<220>
<223> completely synthesized primer
<400> 7

aagtcaaagc aggggttgc 19

<210> 8
<211> 725
<212> PRT
<213> Escherichia coli
<220>
<400> 8

Met Arg Ile Asn Lys Ile Leu Trp Ser Leu Thr Val Leu Leu Val
5 10 15
Gly Leu Asn Ser Gln Val Ser Val Ala Lys Tyr Ser Asp Asp Asp
20 25 30
Asn Asp Glu Thr Leu Val Val Glu Ala Thr Ala Glu Gln Val Leu
35 40 45
Lys Gln Gln Pro Gly Val Ser Val Ile Thr Ser Glu Asp Ile Lys
50 55 60
Lys Thr Pro Pro Val Asn Asp Leu Ser Asp Ile Ile Arg Lys Met
65 70 75
Pro Gly Val Asn Leu Thr Gly Asn Ser Ala Ser Gly Thr Arg Gly
80 85 90
Asn Asn Arg Gln Ile Asp Ile Arg Gly Met Gly Pro Glu Asn Thr
95 100 105
Leu Ile Leu Ile Asp Gly Val Pro Val Thr Ser Arg Asn Ser Val
110 115 120
Tyr Ser Trp Arg Gly Glu Arg Asp Thr Arg Gly Asp Thr Asn Arg
125 130 135
Trp Val Pro Pro Glu Gln Val Glu Arg Ile Glu Val Ile Arg Gly
140 145 150
Pro Ala Ala Ala Arg Tyr Gly Ser Gly Ala Ala Gly Gly Val Val
155 160 165
Asn Ile Ile Thr Lys Arg Pro Thr Asn Asp Trp His Gly Ser Leu
170 175 180
Ser Leu Tyr Thr Asn Gln Pro Glu Ser Ser Glu Glu Gly Ala Thr
190 195 200
Arg Arg Ala Asn Phe Ser Leu Ser Gly Pro Leu Ala Gly Asp Ala
205 210 215
Leu Thr Thr Arg Leu Tyr Gly Asn Leu Asn Lys Thr Asp Ala Asp
220 225 230
Ser Trp Asp Ile Asn Ser Pro Val Gly Thr Lys Asn Ala Ala Gly

```
                    235                     240                       245
His Glu Gly Val Arg Asn Lys Asp Ile Asn Gly Val Val Ser Trp
                245                 250                 255
Lys Leu Asn Pro Gln Gln Ile Leu Asp Phe Glu Val Gly Tyr Ser
                260                 265                 270
Arg Gln Gly Asn Ile Tyr Ala Gly Asp Thr Gln Asn Ser Ser Ser
                275                 280                 285
Ser Ala Val Thr Glu Ser Leu Ala Lys Ser Gly Lys Glu Thr Asn
                290                 295                 300
Arg Leu Tyr Arg Gln Asn Tyr Gly Ile Thr His Asn Gly Ile Trp
                305                 310                 315
Asp Trp Gly Gln Ser Arg Phe Gly Val Tyr Tyr Glu Lys Thr Asn
                320                 325                 330
Asn Thr Arg Met Asn Glu Gly Leu Ser Gly Gly Gly Glu Gly Arg
                335                 340                 345
Ile Leu Ala Gly Glu Lys Phe Thr Thr Asn Arg Leu Ser Ser Trp
                350                 355                 360
Arg Thr Ser Gly Glu Leu Asn Ile Pro Leu Asn Val Met Val Asp
                365                 370                 375
Gln Thr Leu Thr Val Gly Ala Glu Trp Asn Arg Asp Lys Leu Asp
                380                 385                 390
Asp Pro Ser Ser Thr Ser Leu Thr Val Asn Asp Arg Asp Ile Ser
                395                 400                 405
Gly Ile Ser Gly Ser Ala Ala Asp Arg Ser Ser Lys Asn His Ser
                410                 415                 420
Gln Ile Ser Ala Leu Tyr Ile Glu Asp Asn Ile Glu Pro Val Pro
                425                 430                 435
Gly Thr Asn Ile Ile Pro Gly Leu Arg Phe Asp Tyr Leu Ser Asp
                440                 445                 450
Ser Gly Gly Asn Phe Ser Pro Ser Leu Asn Leu Ser Gln Glu Leu
                455                 460                 465
Gly Asp Tyr Phe Lys Val Lys Ala Gly Val Ala Arg Thr Phe Lys
                470                 475                 480
Ala Pro Asn Leu Tyr Gln Ser Ser Glu Gly Tyr Leu Leu Tyr Ser
                485                 490                 495
Lys Gly Asn Gly Cys Pro Lys Asp Ile Thr Ser Gly Gly Cys Tyr
                500                 505                 510
Leu Ile Gly Asn Lys Asp Leu Asp Pro Glu Ile Ser Val Asn Lys
                515                 520                 525
Glu Ile Gly Leu Glu Phe Thr Trp Glu Asp Tyr His Ala Ser Val
                530                 535                 540
Thr Tyr Phe Arg Asn Asp Tyr Gln Asn Lys Ile Val Ala Gly Asp
                545                 550                 555
Asn Val Ile Gly Gln Thr Ala Ser Gly Ala Tyr Ile Leu Lys Trp
                560                 565                 570
Gln Asn Gly Gly Lys Ala Leu Val Asp Gly Ile Glu Ala Ser Met
                575                 580                 585
Ser Phe Pro Leu Val Lys Glu Arg Leu Asn Trp Asn Thr Asn Ala
                590                 595                 600
Thr Trp Met Ile Thr Ser Glu Gln Lys Asp Thr Gly Asn Pro Leu
                605                 610                 615
Ser Val Ile Pro Lys Tyr Thr Ile Asn Asn Ser Leu Asn Trp Thr
                620                 625                 630
Ile Thr Gln Ala Phe Ser Ala Ser Phe Asn Trp Thr Leu Tyr Gly
                635                 640                 645
```

```
Arg Gln Lys Pro Arg Thr His Ala Glu Thr Arg Ser Glu Asp Thr
                650                 655                 660
Gly Gly Leu Ser Gly Lys Glu Leu Gly Ala Tyr Ser Leu Val Gly
                665                 670                 675
Thr Asn Phe Asn Tyr Asp Ile Asn Lys Asn Leu Arg Leu Asn Val
                680                 685                 690
Gly Val Ser Asn Ile Leu Asn Lys Gln Ile Phe Arg Ser Ser Glu
                695                 700                 705
Gly Ala Asn Thr Tyr Asn Glu Pro Gly Arg Ala Tyr Tyr Ala Gly
                710                 715                 720
Val Thr Ala Ser Phe
                725
```

<210> 9
<211> 33
<212> DNA
<213> Artificial Sequence
<220>
<223> forward primer for PCR
<400> 9

cgcgcgcgga tccgacgaga ctctggtggt gga 33

<210> 10
<211> 33
<212> DNA
<213> Artificial Sequence
<220>
<223> reverse primer for PCR
<400> 10

cgcgcgcaag cttgaatgat gcggtaactc cgg 33

# INTERNATIONAL SEARCH REPORT

International application No.
PCT/US00/26117

**A. CLASSIFICATION OF SUBJECT MATTER**

IPC(7) :Please See Extra Sheet.
US CL :536/23.1, 23.7; 435/69.1, 69.3, 71.1, 320.1; 530/350; 424/250.1
According to International Patent Classification (IPC) or to both national classification and IPC

**B. FIELDS SEARCHED**

Minimum documentation searched (classification system followed by classification symbols)

U.S. : 536/23.1, 23.7; 435/69.1, 69.3, 71.1, 320.1; 530/350; 424/250.1

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

WEST, DIALOG, MEDLINE, EMBASE, BIOSIS, CA SEARCH

**C. DOCUMENTS CONSIDERED TO BE RELEVANT**

| Category* | Citation of document, with indication, where appropriate, of the relevant passages | Relevant to claim No. |
|---|---|---|
| X, P<br>-----<br>Y, P | JOHNSON et al. Molecular Epidemiological and Phylogenetic Associations of Two Novel Putative Virulence Genes, iha and ironE.coli' among Escherichia coli Isolates from Patients with Urosepsis. Infect. Immun. May 2000. Vol. 68. No. 5. pages 3040-3047, see entire document. | 1-8<br>-----<br>9-13 |
| A | RUSSO et al. Identification, Genomic Organization, and Anaylsis of the Group III Capsular Polysaccharide Genes kpsD, kpsM, kpsT, and kpsE from an Extrointestinal Isolate of Escherichia coli (CP9, O4/K54/H5). J. Bacteriol. January 1998, Vol. 180. No. 2. pages 338-349, see entire document. | 1-13 |

[X] Further documents are listed in the continuation of Box C. [ ] See patent family annex.

* Special categories of cited documents:

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"&" document member of the same patent family

| Date of the actual completion of the international search | Date of mailing of the international search report |
|---|---|
| 30 NOVEMBER 2000 | 02 FEB 2001 |

| Name and mailing address of the ISA/US | Authorized officer |
|---|---|
| Commissioner of Patents and Trademarks<br>Box PCT<br>Washington, D.C. 20231 | JENNIFER GRASER<br>TERRY J. DEY<br>PARALEGAL SPECIALIST<br>TECHNOLOGY CENTER 1600 |
| Facsimile No. (703) 305-3230 | Telephone No. (703) 308-0196 |

# INTERNATIONAL SEARCH REPORT

International application No.
PCT/US00/26117

C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

| Category* | Citation of document, with indication, where appropriate, of the relevant passages | Relevant to claim No. |
|---|---|---|
| A, P | RUSSO et al. Identification of Genes in an Extraintestinal Isolate of Escherichia coli with Increased Expression after Exposure to Human Urine. Infect. Immun. October 1999, Vol. 67. No. 10. pages 5306-5314, see entire document. | 1-13 |

# INTERNATIONAL SEARCH REPORT

International application No.
PCT/US00/26117

A. CLASSIFICATION OF SUBJECT MATTER:
IPC (7):

C07H 21/02, 21/04; C12P 21/06, 21/04; C12N 15/09, 15/00; C07K 1/00; A61K 39/095

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

CORRECTED VERSION

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
29 March 2001 (29.03.2001)

PCT

(10) International Publication Number
WO 01/021636 A1

(51) International Patent Classification[7]: C07H 21/02, 21/04, C12P 21/06, 21/04, C12N 15/09, 15/00, C07K 1/00, A61K 39/095

(21) International Application Number: PCT/US00/26117

(22) International Filing Date: 22 September 2000 (22.09.2000)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
60/155,621 22 September 1999 (22.09.1999) US

(71) Applicant: THE RESEARCH FOUNDATION OF STATE UNIVERSITY OF NEW YORK [US/US]; State University of New York at Buffalo, Suite 200 UB Commons, 520 Lee Entrance, Amherst, NY 14228 (US).

(72) Inventors: RUSSO, Thomas; 19 Chapel Woods, Amherst, NY 14221 (US). CARLINO, Ulrike; 18 Abbottsford Place, Buffalo, NY 14213 (US).

(74) Agents: KADLE, Ranjana et al.; Hodgson, Russ, Andrews, Woods & Goodyear, LLP, One M & T Plaza, Suite 2000, Buffalo, NY 14203-2391 (US).

(81) Designated States *(national)*: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CR, CU, CZ, DE, DK, DM, DZ, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, UZ, VN, YU, ZA, ZW.

(84) Designated States *(regional)*: ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

**Published:**
— *with international search report*

(48) Date of publication of this corrected version: 14 November 2002

(15) Information about Correction:
see PCT Gazette No. 46/2002 of 14 November 2002, Section II

*For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.*

WO 01/021636 A1

(54) Title: IDENTIFICATION OF A VACCINE CANDIDATE FROM AN EXTRAINTESTINAL ISOLATE OF E. COLI

(57) Abstract: The present invention is directed to identifying a gene, *iroN$_{ec}$*, from an extraintestinal isolate of E. coli, CP9. This gene was identified from CP9 by screening a library of 527 mutant derivatives of CP9 with active TnphoA fusion in human urine. Two mutant derivative CP9.45 and CP9.82 possessed increased PhoA activity in urine due to the Tn*pho*A insertion into *iroN$_{ec}$*. The product of this gene is an extracytoplasmic protein of 725 amino acids and can be used as vaccine against extraintestinal *E. coli* infections.

# IDENTIFICATION OF A VACCINE CANDIDATE FROM AN EXTRAINTESTINAL ISOLATE OF E. COLI

This applications claims the priority of U.S. provisional application serial no. 60/155,621 filed on September 22, 1999.

## FIELD OF THE INVENTION

This invention relates generally to the field of vaccines for bacterial infections. More particularly, this invention identifies a gene that is expressed in extraintestinal isolates of Escherichia coli (*E.coli*), and can be used as an immunogen in vaccine formulations.

## BACKGROUND OF THE INVENTION

*E.coli* is part of the normal intestinal flora where it does not cause infections. However, infections can occur if the bacteria gain entrance to other tissues and organs. This group of *E. coli* strains has been designated herein as extraintestinal pathogenic *Escherichia coli* (ExPEC). Extraintestinal infections (EIs) due to *E. coli* are common in all age groups and can involve nearly any organ or anatomical site. Typical EIs include urinary tract infection (UTI), meningitis (mainly in neonates and following neurosurgery), diverse intra-abdominal infection, pneumonia (particularly in hospitalized and institutionalized patients), intra-vascular device infection, osteomyelitis, and soft tissue infection, which usually occurs in the setting of tissue compromise. Bacteremia can accompany infection at any of these sites.

Extraintestinal isolates of *E. coli* are responsible for the majority of urinary tract infections (UTI). Eighty to ninety percent of ambulatory UTI, 73% in individuals over 50, and 25% of nosocomial UTI are due to extraintestinal strains of E. coli. Uncomplicated urethritis or cystitis occurs most commonly. However,

more severe sequelae of UTI includes pyelonephritis, intrarenal and perinephric abscess, and bacterimia with or without septic shock. Thus, it is clear that despite effective antimicrobial therapy, UTI due to E. coli causes considerable morbidity and mortality.

Although *E. coli* is considered to be a community-acquired pathogen, it also is the most frequently isolated gram-negative bacillus in long-term care facilities and hospitals. Severe illness and death can occur in otherwise healthy hosts, but adverse outcomes are considerably more likely in the presence of comorbid disease and impaired host defenses.

The scope and magnitude of infection caused by extraintestinal strains of *E. coli* is as great as any invasive bacterial pathogen. In fact, perhaps because disease due to extraintestinal isolates of *E. coli* is so common, the virulence and morbidity of this organism is often overlooked. As a result, extraintestinal strains of *E. coli* continue to be low profile "silent killers" imposing a medical-economic strain on the health care system.

Currently, no effective vaccine is available against ExPEC. To date, efforts to identify specific vaccine candidates against ExPEC have concentrated on virulence traits such as capsule, LPS, and pili. These studies have demonstrated that antibodies directed against these structures confer protection against homologous strains *in vivo*. However, the inherent marked antigenic variability of these components may limit their utility as vaccine candidates. Thus, there is an ongoing need for identification of novel effective strategies for the prevention and treatment of ExPEC infections including UTI.

Summary Of the Invention

The present invention is directed to identifying a gene, *iroN*$_{ec}$, from an extraintestinal isolate of *E. coli*. This gene was identified by transposon mutagenesis. Tn*phoA* mutagenesis was performed on the wild type isolate of *E.coli*, CP9, and a library of 527 mutants with active Tn*phoA* fusions was generated and saved. This library was subsequently screened for mutant derivatives with increased PhoA activity in the presence of urine. By this method, the *iroN*$_{ec}$ gene was identified which encodes a protein of 725 amino acids. This protein has an extracytoplasmic location. DNA homology data and data demonstrating that its transcription is iron repressed supports its function as a siderophore receptor. High stringency Southern hybridization identified DNA sequences homologous to *iroN*$_{ec}$ in 80-93% of ExPEC strains.

This protein or antigenic epitopes thereof can be used for inducing an immune response against extraintestinal *E.coli* infections. Data is presented to show that IroN$_{ec}$ is strongly immunogenic (without adjuvant) in mice. Further, antibodies developed against IroN$_{ec}$ were observed to be protective in a mouse intra-peritoneal challenge model. Mice immunized with IroN$_{ec}$ had diminished mortality after intra-peritoneal challenge with the *E. coli* strain CP9. Additionally, there was diminished growth of the challenged strain in the liver and spleen in animals immunized with IroN$_{ec}$ compared to non-immunized controls.

Thus, it is an object of the present invention to identify a gene, designated herein as *iroN*$_{ec}$ in the extraintestinal isolates of *E.coli* with increased expression in human urine.

It is another object of the present invention to provide a polynucleotide that encodes the protein, $IroN_{ec}$.

It is another object of the present invention to provide polynucleotides that hybridize, preferably under high stringency conditions, with a polynucleotide encoding $IroN_{ec}$.

It is anther object of the present invention to provide peptides that are encoded by $iroN_{ec}$ and other polynucleotides of the present invention.

It is another object of the invention to provide antigenic compositions comprising the peptides or antigenic fragments thereof for the treatment or prevention of extraintestinal infections caused by ExPEC.

## BRIEF DESCRIPTION OF THE DRAWINGS

Figure 1 is a schematic diagram of the $iroN_{ec}$ and the flanking sequence.

Figure 2A is a representation of the SDS-PAGE analysis of $IroN_{ec}$ purification. Lane 12 represents molecular weight markers. The approximate molecular weight of each marker is indicated. Lane 11 is non-purified sample of the induced culture and lanes 1-10 are eluted $IroN_{ec}$.

Figure 2B is a representation of Western blot analysis of SDS-PAGE from Figure 2A. An antibody to the recombinant $IroN_{ec}$ was used for detection. Lane 12 represents molecular weight markers having the indicated approximate molecular weights.

## DESCRIPTION OF THE PREFERRED EMBODIMENTS

The present invention is directed to a gene whose expression is increased in human body fluids such as

urine, blood and ascites. This gene, termed *iroN*$_{ec}$, is represented by the sequence of SEQ ID NO:1.

In one aspect of this invention are provided polynucleotides that are at least 90% identical (homologous) over their entire length to the polynucleotide of SEQ ID NO:1, preferably at least 95% identical, and still more preferably at least 97% identical to the sequence of SEQ ID NO:1, or complementary sequences thereof.

The invention further relates to polynucleotides that hybridize to the sequence of SEQ ID NO:1 under high stringency hybridization conditions. High stringency conditions as used herein means hybridization will occur only if there is at least 90% homology, preferably at least 95% homology, and even more preferably at least 97-99% homology between the sequences. An example of high stringency hybridization conditions is overnight incubation at 42°C in 50% formamide, 5 X SSC, 50 mM sodium phosphate (pH 7.6), 5X Denhardt's solution, 10% dextran sulfate, and 20 μg/ml denatured, sheared salmon sperm DNA. Washing is carried out in 0.1 X SSC at about 65°C. Details on the stringent hybridization conditions are well known to those skilled in the art (see Sambrook et al., 1989, Molecular Cloning: A laboratory Manual, Second Edition, Cold Spring Harbor, NY). Another example of high stringency hybridization condition is presented in Example 2.

The polynucleotides of the present invention can be incorporated into vectors which can be used in expression systems for the production of polypeptides. A variety of expression systems are available and known to those skilled in the art.

In another embodiment of the invention are provided peptides that are encoded by the polynucleotides of the present invention. These peptides or antigenic

fragments thereof, can be used for diagnostic and immunotherapeutic purposes. "Antigen fragments" are those that are part of the whole protein and are specifically recognized by certain antibodies that also recognize the whole protein and which will generate an immune response that reduces or prevents the incidence or symptoms of *E.coli* infections.

In the present invention, $iroN_{ec}$ was identified by transposon mutagenesis. Transposons or transposable elements are DNA segments that can move in the genome from one site to another. They can also be moved from one bacterium to another by plasmid or other vehicles to be inserted into the host genome. Transposons may be simple or complex. Simple transposons have insertion sequences and no genes other than those involved in their own transposition into the target sequence. Complex transposons contain the insertion elements bracketing additional genes that encode properties such as drug resistance, carbohydrate metabolism, light generation, ice nucleation, and other properties which can function as selectable or screenable markers for the entire transposable elements. Thus, transposons are suitable for the use of identifying and cloning of bacterial genes which are turned off or on in the presence of a particular environment.

The method used for the identification of this gene has been described in Russo et al. (1996, Mol. Microbiology, 22:217-229), the disclosure of which is incorporated herein by reference. A transposon, Tn*phoA*, which is a Tn5 derivative, was used in the present invention. On correct insertion into an open reading frame, this transposon will produce a fusion protein with bacterial alkaline phosphatase (PhoA). Since PhoA is only active when in the periplasm, Tn*phoA* identifies genes whose products are exported across the

cell membrane. When a PhoA-specific chromogenic detection reagent (for example, XP; 5-bromo-4-chloro-3-indolyl phosphate) is used, a blue colony is only obtained when *TnphoA* has inserted into a gene encoding a protein containing the N-terminal export leader peptide sequence.

The pathogen CP9 (Russo et al., 1993, *Mol. Microbiol.*, 9:357-364), an *E. coli* blood isolate cultured from a patient with sepsis, hospitalized at the National Institutes of Health, was used as a model pathogen for identification of genes with increased expression in urine. It is characterized by growth in 80% normal human serum, ß-hemolysis, no known antibiotic resistance, O4/K54/H5 serotype, P pilus (class I PapG adhesin), Prs pilus (class III PapG adhesin), type 1 pilus, possession of a 36.2-kb cryptic plasmid (pJEG) and an aerobactin minus genotype. By DNA dot-blot assays, it has also been determined to be *sfa*, *ompT*, *cnf1* and *drb* positive (Johnson et al., 1997, *Infect. Immun.*, 65:2153-2159; Russo et al., 1993, *Mol. Microbiol.*, 9:357-364). It is highly virulent in a mouse UTI infection model (Russo et al., 1996, *Infec. Immun.*, 64:2343-2348). Recent studies have established that CP9 is part of a widely disseminated group of uropathogens that are characterized, in part, by possessing group 3 capsules, the O4 specific antigen, and Class 1 and 3 Pap adhesins (Johnson et al., 1997, *Infect. Immun.*, 65:2153-2159; Johnson et al., 1997, *Infect. Immun.*, 65:2153-2159).

All strains of *E.coli* were maintained at -80°C in 50% Luria-Bertani (L-B) medium and 50% glycerol. L-B broth consisted of 5 grams yeast extract, 10 grams tryptone, 10 grams NaCl per liter (L). Incubations were performed at 37°C unless otherwise described. For

plates, 15 grams of Bacto-Agar (Difco Laboratories , Detroit, MI) were added per L and kanamycin (kan) (40 $\mu$g/mL) or ampicillin (200 $\mu$g/ml) (Amresco, Solon, OH) were added where appropriate. Urine agar plates were made as described, using pooled urine from 5 healthy donors that did not have a history of having a urinary tract infection (Russo et al., 1996, *Mol. Microbiol.*, 22:217-229).

**EXAMPLE 1**

This embodiment describes the identification, cloning, and characterization of *iroN*$_{ec}$.

To identify genes with increased expression in human urine *ex vivo* that coded for extracytoplasmically located gene products, random Tn*phoA* mutagenesis was performed on CP9. A Tn*phoA* mutant library, consisting of 527 CP9 derivatives containing active Tn*phoA* fusions had been previously constructed (Russo et al., 1993, *Mol. Microbiol.*, 9:357-364) and was used in the present study.

The Tn*phoA* mutant library was screened to identify genes that coded for extra-cytoplasmically located gene products which had increased expression in human urine *ex vivo*. Transposon mutants were intially plated on human urine agar plates containing kanamycin (40$\mu$g/ml) and 5-bromo-4-chloro-3-indolyl-phosphate-p-toluidine salt (XP), as the colorimetric substrate for an indication of alkaline phosphatase activity. Mutants that appeared to have increased expression via these qualitative screens (as manifested by the relative blue color of the colonies) were confirmed with quantitative assays as follows. For urine and L-B assays quantitative alkaline phosphatase assays were performed as previously described (Russo et al., 1996, supra, Russo et al.,

1993, supra). In brief, CP9 and each of its mutant derivatives were grown overnight in L-B broth and human urine. Preliminary experiments established that PhoA expression was the same for CP9 derivatives containing an active PhoA fusion in *iroN*$_{ec}$ (CP9.45, CP9.82) in log phase and stationary phase grown cells. Cells were then washed, permeabilized, and p-nitrophenyl phosphate was added for detection of alkaline phosphatase activity respectively. To control for both endogenous bacterial alkaline phosphatase activity and any activity from the growth medium (e.g. urine) that may persist despite washing, alkaline phosphatase activity from CP9 was subtracted from the measured activity of its mutant derivatives.

The screening of the Tn*phoA* library resulted in the identification of CP9.45 and CP9.82 (Table 1).

Table 1

*E. coli* strains and plasmids used in this study

| Strain/ Plasmid | Genotype or other relevant characteristics | Derivation |
|---|---|---|
| Strains | | |
| CP9 | 04/K54/H5 | clinical blood isolate |
| CP9.45 | *iroN*$_{ec}$:Tn*phoA*, active Tn*phoA* fusion | Kan$^R$exconjugant of CP9/pRT291 x MM294/pPH1J1 |
| CP45 | *iroN*$_{ec}$:Tn*phoA* | T4(CP9.45) x CP9 |
| CP9.82 | *iroN*$_{ec}$:Tn*phoA*, active Tn*phoA* fusion | Kan$^R$exconjugant of CP9/pRT291 x MM294/pPH1J1 |
| CP82 | *iroN*$_{ec}$:Tn*phoA* | T4(CP9.82) x CP9 |

| Plasmid | | |
|---|---|---|
| p45.1 | 21kb*Sal1/SacI* fragment from CP45 containing the leftward 5.0kb of Tn*phoA* (active fusion) and a portion of *iroN*$_{ec}$ cloned into pBS II SK (-) | |
| p45.3 | 22kb *Cla1/Xba1* fragment from CP45 containing the rightward 6.7kb of Tn*phoA* (non-fusion) and a portion of *iroN*$_{ec}$ cloned into pBS II SK (-) | |
| p82.1 | 22kb *BamH1/Sac1* fragment from CP82 containing the leftward 5.0kb of Tn*phoA* (active fusion) and a portion of *iron*$_{ec}$ cloned into pBS II SK (-) | |

To establish that these mutants were isogenic derivatives of CP9, a series of studies were performed to exclude the possibility that cryptic mutations were acquired during the mutagenesis procedure.

Transduction of the transposon insertion back into the wild-type strain CP9 was accomplished using the bacteriophage T4 resulting in transductants CP45 and CP82. Whole cell DNA preparation, restriction enzyme (New England Biolabs, Beverly, MA.) mediated DNA digestion and Southern hybridization using PCR generated radioactive probes was performed as described (Russo et

al., 1993, supra). Primers 63 (SEQ ID NO:2) and 64 (SEQ ID NO:3) were used to amplify a 4.0kb internal fragment of Tn*phoA* (contained in pRT291), which were used to probe for Tn*phoA* insertions. Southern analysis of *Bgl II* digested whole cell DNA containing Tn*phoA* produces a 2.8kb internal fragment and two variable junctions fragment per copy. Lipopolysaccharide (LPS), capsular polysaccharide, and outer membrane protein profiles were determined as described (Russo et al., 1995, *Infec. Immun.*, 63:1263-1269).

In the mutants CP9.45 and CP9.82, and their respective T4 generated transductants, CP45 and CP82, it was established that the mutants had a single transposon insertion, that transduction of this insertion back into the wild type strain (CP9) resulted in a derivative that possessed the same degree of increased expression in urine relative to L-B medium as the original mutant, that in the transductants the transposon was physically in the same location as in the original mutant and that alterations in capsule, lipopolysaccharides and outer membrane profiles did not occur. This data establishes that CP45 and CP82 are isogenic derivatives of CP9 and that the gene into which TnphoA was inserted ($iroN_{ec}$) has increased expression in urine.

Subclones of the gene loci 5' to the Tn*phoA* insertions in CP9.45($iroN_{ec}$) were obtained by restricting whole cell DNA with *BamHI* or *SalI*, which recognizes a site located 3' to the kanamycin resistance gene in Tn*phoA* plus *SacI* (CP45, CP82) which does not have restriction sites within Tn*phoA*. Ligation of this restriction into pBSII SK(_), electroporation into XL1 Blue (Stratagene, Lajolla, CA) and selection of ampicillin (200$\mu$g/ml) and kanamycin (40$\mu$g/ml) resistant transformants resulted in the identification of the

subclones p45.1 and p82.1 (Table 1). A subclone of the *iroN*$_{ec}$ gene locus 3' to the Tn*phoA* insertion in CP9.45 was obtained by restricting whole cell DNA with *ClaI* which recognizes a site 5' to the kanamycin resistance gene in Tn*phoA* and *XbaI* which does not possess a restriction site within Tn*phoA*. Ligation of this restriction into pBSII SK(-), electroporation into XL1 Blue (Stratagene, Lajolla, CA), and selection of ampicillin and kanamycin resistant transformants resulted in the identification of the subclone p45.3.

DNA sequence was determined by the dideoxy chain termination method of Sanger (Sanger, et al, 1977, *Proc. Natl. Acad. Sci*. 85:5463-5467) using the gene subclones p45.1, p45.3 and p82.1 as the DNA templates. DNA sequencing of the gene subclones p45.1 and p82.1 initially utilized a Tn*phoA* fusion joint primer (SEQ ID NO:4) which established the location for a given Tn*phoA* insertion. Sequencing of the gene subclone p45.3 initially utilized the Tn*phoA* primer (SEQ ID NO:5). Subsequent DNA sequence was determined using primers derived from the deduced sequences of the gene subclones. A consensus sequence for *iroN*$_{ec}$ was generated by assembling and editing the DNA sequence obtained from 34 overlapping but independent sequencing reactions using AssemblyLIGN 1.0.2 (Oxford Molecular Group, Beaverton, OR). Both strands of the gene were sequenced. The sequence of *iroN*$_{ec}$ is disclosed in SEQ ID NO:1 (accession no. AF135597). Sequence analysis, comparisons, and CLUSTAL alignments were performed, in part, using MacVector (version 6.0, Oxford Molecular Group, Beaverton, OR). Comparisons were also performed via BLAST analysis of the non-redundant GenBank+EMBL+DDBJ+PDB sequences. SignalP V1.1 was used for identification of signal sequences (Perez-Trallero

et al., 1993, *Eur. J. Clin. Microbiol. Infect. Dis.*, 12:349-351).

The DNA sequence indicates that this is a novel gene for *E. coli*. The predicted protein ($IroN_{ec}$) consists of 725 amino-acids, a putative molecular weight of 79,380 kDa, an estimated pI of 5.68, and the first 24 a.a. represent a putative signal sequence. Comparison with entries in GenBank revealed that it was most homologous with $iroN_{se}$ (77% nucleotide homology), a recently identified catecholate siderophore receptor in *Salmonella enterica*. (Baumler, et al, 1998, 180:1446-1453). In *S. enterica*, $iroN_{se}$ is part of a five gene cluster that includes *iroBCDEN*. The specific functions of the gene products for *iroBCDE* remains unclear and a gene encoding the cognate siderophore for $IroN_{se}$ has yet to be identified. However, a putative Fur DNA binding site is present 5' to both $iroN_{se}$ (86 bases) and $iroN_{ec}$ (100 bases) with 15/19 bases conserved. Further, 184 bases 5' to the start site of $iroN_{ec}$ are 137 bases which were 87% (120/137) homologous to the *Salmonella* insertion sequence IS*1230* (bases 1-137, accession no. AJ000635). These bases correspond to the first 137 bases from this IS3-like element, the first 39 of which consist of an imperfect inverted repeat. (Collighan, et al, 1997, *FEMS Micobiol. Lett.* 154:207-213). This IS element was not present 5' to $iroN_{se}$. Taken together, this data suggests that $iroN_{ec}$ may have been acquired from *S. enterica* via IS-element mediated horizontal transfer. $iroN_{ec}$ possessed a lesser degree of homology with the catecholate siderophore receptors from *Pseudomonas aeruginosa* (*pfeA*), *Bordetella pertussis* (*bfeA*), and the *fepA* siderophore receptor from *E. coli*.

The deduced protein sequence identity and similarity of $IroN_{ec}$ determined by pairwise amino acid alignment using the program CLUSTAL to $IroN_{se}$, FepA, PfeA, and BfeA is summarized in Table 2. Further, sequence analysis did not reveal any premature stop codons.

Table 2

Deduced protein sequence identity and similarity of $IroN_{ec}$ to $IroN_{se}$ FepA, and BfeA

| | Homologue | % Identity | % Similarity |
|---|---|---|---|
| *Salmonella enterica* | $IroN_{ec}$ | 82% | 91% |
| *Escherichia coli* | FepA | 52% | 69% |
| *Pseudomonas aeruginosa* | PfeA | 52% | 69% |
| *Bordetella pertussis* | BfeA | 53% | 68% |

Additional sequence analysis of DNA 5' to *iroN*$_{ec}$ has demonstrated that this gene is 1.6kb 3' to the *prs* gene cluster, which encodes the class III PapG adhesin (Figure 1). Further, although the precise genomic organization of the region 5' to the *prs* operon has not been determined, the molecular usher for the F1C fimbria ((Klemm, et al, 1995, *J. Bacteriol*. 177:621-627) *focD*, has also been identified approximately 15.6kb 5' to the deduced start site for *iroN*$_{ec}$ (Figure 1). From left to right in figure 1 are shown: 11.8 kb sequence represented by the dotted line consists of the 5' portion of the *iroD*$_{ec}$ gene and genome 5' to it; a portion of the *iroD*$_{ec}$ gene (0.6kb), *iroE*$_{ec}$ (0.9kb) and

*iroN$_{ec}$*(2.2kb); a putative Fur binding site which is 100 bases 5' to the deduced start site of iroN$_{ec}$; a portion of an IS12230-like element (bp 1-137, accession number AJ000635) which is 184 bp 5' to the deduced start site of *iroN$_{ec}$*; and a 14 kb sequence represented by the dotted line (not to scale relative to solid line). The most distal boundary of this sequence consists of the 3' portion of focD, the molecular usher for the FIC fimbria. The arrow below the solid line indicate the direction of transcription.

**Example 2**

This embodiment describes the phylogenetic distribution of *iroN$_{ec}$* amongst various isolates of *E. coli*. A 667 base pair internal DNA probe (AF135597 bases 1729-2396) was generated from *iroN$_{ec}$* which did not share any homology with *fepA*. This probe was used in a dot-blot assay as described below to detect for the presence of homologous *iroN$_{ec}$* sequence.

DNA was prepared from relevant strains by boiling cells from overnight growth in L-B medium (1ml concentrated to 200μl of sterile $H_2O$) at 105°C for 10 minutes. The supernatant was saved and used for analysis. Nytran membranes (Schleicher & Scheull, Keene, New Hampshire) were pre-wet in 6X SSC for 10 minutes and dotted with 3 μl of denatured DNA preparation from each strain in triplicate. The membrane was subsequently placed on filter paper saturated with denaturing solution (0.4N NaOH, 0.6M NaCl) followed by neutralizing solution (1.5M NaCl, 0.5M Tris HCl, pH 7) for 1 minute each, then air-dried, and UV cross-linked with 1200 Joules (UV Stratalinker 2400, Statagene, La Jolla, CA). Primers 192 (SEQ ID NO:6) and 197 (SEQ ID NO:7) were used to amplify a 0.67kb internal fragment of *iroN$_{ec}$*

(contained in p45.3) that did not share any homology with *fepA*. Aqueous hybridization was performed under high stringency conditions (65°C). Results were scored as either zero (no hybridization), $1^+$, or $2^+$. Under these conditions the negative control strains (HB101, XL-1 Blue) were consistently scored as zero and the positive control strain CP9 as $2^+$. Experimental strains consisted of; Group 1: fourteen unique fecal isolates that had been previously established not to contain *pap*, *hly*, or *cnf-1* (Johnson, et al, 1998, *J. Infect. Dis.* 177:1120-1124), Group 2: five unique fecal isolates that possessed some combination of *pap*, *hly*, or *cnf-1*, Group 3: twenty unique first-time UTI isolates (Russo, et al, 1995, *J. Infect. Dis.* 172:440-445), Group 4: fifteen unique recurrent UTI isolates (Russo, et al, 1995, *J. Infect. Dis.* 172:440-445), Group 5: twenty-one blood isolates obtained from patients hospitalized at Erie County Medical Center (Buffalo, NY), Group 6: all 35 UTI isolates, and Group 7: all 56 clinical isolates. Group 1 was most representative of non-pathogenic or commensal strains and therefore was used in statistical comparisons against the clinical isolate Groups 5, 6, and 7. Fisher's exact test was used for the comparison of fecal versus clinical isolates for the presence of $iroN_{ec}$ DNA sequence via dot-blot assay.

The results from these studies are summarized in Table 3.

Table 3

Phylogenetic distribution of *ironN*$_{ec}$ homologous DNA sequence

| Dot-Blot Score | Group 1 | Group 2 | Group 3 | Group 4 | Group 5 | Group 6 | Group 7 |
|---|---|---|---|---|---|---|---|
| 0 | 43% | 20% | 5% | 7% | 10% | 6% | 7% |
| $1^+$ | 43% | 0% | 50% | 27% | 57% | 40% | 46% |
| $2^+$ | 14% | 80% | 45% | 67% | 33% | 54% | 46% |
| any¬ | 57% | 80% | 95% | 93% | 90% | 94% | 93% |

Dot-Blot Score: 0= no homology, $1^+$ = positive homology, $2^+$ = maximal homology, any$^+$ = the combination of $1^+$ and $2^+$

Group 1: fourteen unique fecal isolates that had been previously established not to contain *pap*, *hly*, or *cnf-1* and therefore are most representative of non-pathogenic strains.

Group 2: five unique fecal isolates that possessed some combination of *pap*, *hly*, or *cnf-1* and therefore most likely represent pathogenic strains

Group 3: twenty unique first-time UTI isolates

Group 4: fifteen unique recurrent UTI isolates

Group 5: twenty-one blood isolates obtained from patients hospitalized at Erie County Medical Center (Buffalo, NY).

Group 6: All 35 UTI isolates (Groups 3 and 4).

Group 7: All 56 clinical isolates (Groups 3, 4 and 5).

[a]Fisher's exact test was used for proportions. All comparisons are versus Group 1.

[b]P=0.039, [c]P=0.004, [d]P=0.003, [c]P=>0.10(NS), [f]P=0.01, [g]P=0.03

In summary, forty three percent of 14 fecal isolates (Group1, negative for *pap*, *hly*, or *cnf1* and

therefore most representative of non-pathogenic strains) did not possess DNA sequence homologous to *iroN* $_{ec}$. In contrast, only 20% of 5 fecal isolates (Group 2, positive for some combination of *pap*, *hly*, or *cnf1* and therefore most likely pathogenic strains), 5% of 20 first time UTI isolates (Group 3), 7% of 15 recurrent UTI isolates (Group 4), and 10% of 21 blood isolates (Group 5) were negative for *iroN*$_{ec}$ homologous sequence under high stringency conditions. The differences between Group 1 versus either Group 5, Group 6 (all UTI strains, Groups 3,4), or Group 7 (all clinical isolates, Groups 3,4,5) were statistically significant (P = 0.039, P = 0.004, P = 0.003 respectively). In summary, this data demonstrates that DNA sequence homologous to *iroN*$_{ec}$ is significantly less prevalent in fecal isolates without the virulence genes *pap*, *hly*, or *cnf1* than clinical isolates.

**Example 3**

This embodiment describes the growth of CP9 and CP82 and the expression of *iroN*$_{ec}$ in human urine, human ascites and blood.

<u>Ex vivo growth in human urine</u>

Human urine from subjects who had and who never had experienced a UTI was used for studies assessing growth of strains *ex vivo*. The strain to be tested was grown overnight in 2 ml of L-B medium ± kanamycin 40μg/ml. The next day, the bacterial cells were diluted into urine to achieve a starting concentration of approximately 1.0 x $10^{2-3}$ cfu/ml, since this titer is at the lower end of the spectrum for what is considered significant for UTI in symptomatic young women (Stamm, et al, 1982, *N. Engl. J. Med.* 307:463-468). For $Å_{600}$ growth curves, a starting $Å_{600}$ of about 0.03 was used. During incubation at 37°C,

aliquots were removed at intervals and either the $Å_{600}$ was determined or the bacterial titers were established by plating 10-fold serial dilutions in 1X phosphate-buffered saline in duplicate on appropriate media.

Expression of $iroN_{ec}$ in urine

Expression of *iroN*$_{ec}$ was increased in human urine relative to L-B. Since the composition of human urine has the potential to be variable, assays were performed using 17-29 independent urines collected from 10 different individuals. Five of these individuals were women with a prior history of UTIs. For CP9.82, the mean fold and median fold increase in expression of $IroN_{ec}$ was $27.2 \pm 5.0$ and 19.0 respectively, the range being 2.4-132. Although there was variance in the degree of increased expression from urine to urine, increased expression was seen in all urines evaluated. The degree of expression of *iroN*$_{ec}$ was similar in urines from individuals with and without a prior history of UTI. The 17-29 independent urines used were filter sterilized and stored at 4°C prior to use. To determine if the processing of urine affected gene expression, assays were performed in parallel using urines that were either 1) fresh and unfiltered, 2) fresh and filtered with a .22 micron filter, or 3) filtered and stored at 4°C. The expression of *iroN*$_{ec}$ was similar, regardless of how the urine was processed (data not shown).

Growth of CP9 (w.t.) and CP82 (*iroN*$_{ec}$) was evaluated in multiple independent urines via both enumeration of bacterial titers and $Å_{600}$. Growth of CP82 (5 urines) was equivalent to their wild-type parent CP9 (data not shown).

Expression of *iroN*$_{ec}$ in human ascites and blood

The expression of various virulence traits may vary depending on the site of infection. Therefore *iroN*$_{ec}$, expression was evaluated in human blood and ascites, two

additional body fluids which extraintestinal *E. coli* isolates commonly infect. To assess gene expression in human ascites and blood, filter sterilized ascites (peritoneal fluid) was obtained from a patient hospitalized at Erie County Medical Center, divided into multiple aliquots, and frozen at -80°C. Blood was used fresh and was obtained from a single donor. It was collected in sterile, 8.3ml vacutainer tubes which contained 1.7ml of sodium polyanetholesulfonate (0.35%) and NaCl (0.85%) (non-bactericidal) as the anti-coagulant. For blood assays the bacterial cells were washed x 2 at 4°C with 4ml of 0.1M Tris (pH 9.8), 0.001M $MgCl_2$ buffer and resuspended in a total volume of 2ml. For ascites assays the bacterial cells were concentrated via centrifugation and the resultant pellet was resuspended in 2ml of 0.1M Tris (pH 9.8), 0.001M $MgCl_2$ buffer. Aliquots were removed x 2 and cfu/ml were determined via serial 10-fold dilutions. Bacterial cells were subsequently permeabilized by adding 100$\mu$l of 0.1% SDS and 200$\mu$l of chloroform, vortexed x 10 seconds and kept on ice. A fluorescent assay was performed because red blood cells ± hemoglobin present in blood could not be reliably separated from bacterial cells. Their presence interfered with the colorimetric assays described above for measuring alkaline phosphatase activity. Assays were performed in a 48-well tissue culture plate. Each assay mixture consisted of 1ml of Tris buffer, 50$\mu$l of bacterial cell extract, and 50$\mu$l (0.01M) of the fluorescent substrate (4-methylumbelliferonephosphate). Samples were read using a fluorescence multi-well plate reader (CytoFluor II, PerSeptive Biosystems, Framingham, MA) at an excitation setting of 360 nm, an emission setting of 460 nm, and a gain of 80 for 15 cycles. The net sample rate in blood or ascites relative to that in L-B broth established the fold induction. The net sample rate / ml (SR) =

{((fluorescence cycle B - fluorescence cycle A over the linear portion of the curve)/(elapsed time)) x 20} - (CP9 SR). Specific activities were determined by dividing net sample rates by cfu/ml. The sensitivity of the colorimetric and fluorescent assays was established to be similar.

The expression of *iroN*$_{ec}$ in blood and ascites was as follows. The mean fold increase in blood was 65.8±6.7 while the mean fold increase in ascites was 207±27. Although there was some variance in expression compared to urine, it should be noted that ascites and blood were obtained from single individuals.

**EXAMPLE 4**

This embodiment describes the regulation of the expression of *iroN*$_{ec}$ under various environmental conditions.

For osmoregulation studies, modified Davis medium was used with variable concentrations of either NaCl (0.05M-0.7M) or urea (0.05M-0.7M). Some gene regulation studies utilized urine to which exogenous Fe (0.1mM) or glucose (0.5%) was added. M9 minimal medium was also utilized in gene regulation studies. Fe was chelated from M9 medium by mixing 200ml of medium with 21.2 grams of washed (with 1L d$H_2O$ x 2) iminodiacetic acid (Chelex 100, Sigma, St. Louis, MO.) for 90 minutes followed by filter sterilization. Siderophore production was determined using the Arnow assay as described (Schwyn, et al, 1987, *Anal. Biochem.* 160:47-56) and was concomitantly measured to confirm that the Fe concentration was limiting when chelated. As expected, siderophore production increased from 3.1$\mu$M/Å$_{600}$ in the presence of Fe (0.1mM) to 10.7$\mu$M/Å$_{600}$ when Fe was chelated. The effect of pH on expression was determined using pooled urine whose pH was adjusted with either HCl

or NaOH to achieve pHs of 5.0, 6.0, and 7.0. For a given experiment, assays were performed in triplicate and experiments were repeated at least once. Results were presented as the ratio of reporter gene expression in urine relative to L-B medium. For all of these studies quantitative alkaline phosphatase assays were performed as described above.

<u>Iron:</u> Sequence analysis of *iroN*$_{ec}$ strongly supports that this gene codes for a catecholate siderophore receptor, and therefore the role of Fe in the regulation of *iroN*$_{ec}$ was evaluated. Expression of *iroN*$_{ec}$ was measured when CP82 (*iroN*$_{ec}$) was grown in M9 minimal media in which Fe was either chelated or added exogenously. Compared to M9 medium plus Fe, *iroN*$_{ec}$ expression was increased 20.8-fold when CP82 was grown in Fe chelated M9 medium. Further, the addition of exogenous Fe to 3 independent human urines suppressed the increased expression of *iroN*$_{ec}$ relative to L-B medium (mean 64-fold decrease in *phoA* activity). These experimental findings, in combination with the identification of a Fur binding sequence 5' to the start of *iroN*$_{ec}$, suggest that *iroN*$_{ec}$ is Fur- regulated and that Fe is limited in urine.

<u>pH:</u> Although the pH of normal urine most commonly ranges from 5.5 - 6.5, values from 5.0 to 8.0 can occur. Therefore the effect of pH 5.0, 6.0, and 7.0 on the expression of *iroN*$_{ec}$ in human urine was evaluated. The expression of *iroN*$_{ec}$ was completely suppressed at pH 5.0 but unaffected at pH 6.0 and 7.0 with induction ratios of 0.23, 32, and 34 measured respectively. Therefore, urinary pH can affect gene expression of *iroN*$_{ec}$.

Thus, low Fe concentrations increase the expression of *iroN*$_{ec}$, and its expression is suppressed at a urinary pH 5.0, but unaffected by limiting concentrations of amino acids, nucleotides, or glucose.

**EXAMPLE 5**

This embodiment illustrates the use of the $IroN_{ec}$ for antigenic purposes. The entire protein encoded by *$iroN_{ec}$* or antigenic fragments thereof can be used in vaccine formulations. Surface exposed epitopes of $IroN_{ec}$ can be identified by methods well known to those skilled in the art. For vaccine development, the $IroN_{ec}$ protein may be purified from the bacteria or may be purified from host containing a recombinant vector which expresses *$iroN_{ec}$*. The antigenic formulation may be introduced into the human or animal to be vaccinated by standard techniques well known to those skilled in the art.

To illustrate this embodiment, the $IroN_{ec}$ was purified following the cloning of the gene into an expression vector. The purified protein was used to elicit antibodies in mice. Immunized mice were challenged by the homologous strain of *E.coli* and protective abilities of this protocol determined.

<u>Cloning of the *$iroN_{ec}$*</u>

For PCR-mediated amplification, the following 2 primers were designed for the entire *$iroN_{ec}$* gene, excepting its signal sequence (2083 base pairs). These primers were derived from the *iroN* sequence (SEQ ID NO:1). The forward primer CGCGCGCGGATCCGACGAGACTCTGGTGGTGGA (SEQ ID NO:9) and the return primer CGCGCGCAAGCTTGAATGATGCGGTAACTCCGG (SEQ ID NO:10) were used. A single band of the expected size was PCR amplified from CP9 chromosomal DNA. The DNA was cleaned and ligated into the Kanamycin resistant pet28a T7/his-tag expression vector. The pet28a::*iroN* construct was electroporated into XL1 Blue cells and selected for on LB plates containing Kanamycin. The *$iroN_{ec}$* gene in the selected clone was confirmed to be

correct by DNA sequencing. This excluded the possibility that an error was introduced into the cloned *iroN*$_{ec}$ during PCR amplification. The clone was subsequently electroporated into the expression cell line AD494 (DE3) pLysS for over-expression of IroN$_{ec}$.

AD494 (DE3) pLysS::pet28a::*iroN*$_{ec}$ was grown overnight in LB media plus Kanamycin. The next morning, 1 ml of the overnight culture was transferred into 11ml LB media plus Kanamycin and grown at 37°C for 2.5 hours, shaking. IPTG was added to a final concentration of 1mM to induce the expression of IroN$_{ec}$. One ml aliquots of the induced culture as well as an uninduced control culture were taken in 30 minute intervals. The samples were prepared for gel electrophoresis and run on a 7.5% SDS-PAGE gel (Figure 2A). This figure shows the increased expression of IroN$_{ec}$ in the induced culture, migrating at approximately 80 kDa . This size is close to the deduced size of IroN$_{ec}$ (79.4 kDa) based on its DNA sequence. IroN$_{ec}$ was subsequently purified using TALON cobalt-based Immobilized Metal Affinity Chromatography and eluted under denaturing conditions using a 6M urea elution buffer at a pH between 5.1 and 5.3. A Western blot using the T7-Tag antibody specific to the recombinant protein was done. An intact IroN$_{ec}$ is represented by the primary band (Figure 2B). The minor bands below IroN$_{ec}$ represent break-down products of IroN$_{ec}$ since these bands are all recognized by the T7-tag antibody which is specific to recombinant IroN$_{ec}$. This established that IroN$_{ec}$ was successfully purified. Purified IroN$_{ec}$ was stored at -20°C until further use.

The successful expression of IroN$_{ec}$ confirmed that *iroN*$_{ec}$ encoded as a full length product. Further, the over-expressed and purified IroN$_{ec}$ was used as an immunogen in mice to determine if it was antigenic and elicited an antibody response.

Purified $IroN_{ec}$ elicits a strong antibody response in mice

A putative vaccine candidate must be immunogenic. To establish that this is the case for $IroN_{ec}$, mice were immunized with purified $IroN_{ec}$ and sera pre and post-immunization were obtained and tested as follows.

**a) Immunization protocol:**

Balb C mice were used in an immunization protocol that was performed over a 5 week period. Mice were divided into 3 groups: Group 1, controls immunized with buffer only (N=15); Group 2, animals immunized with a total of 70μg of $IroN_{ec}$ (N=12); and Group 3, animals immunized with a total of 150μg of $IroN_{ec}$ (N=15). Purified $IroN_{ec}$ was injected subcutaneously on days 1, 15 and 30 in a total of 200μl for the first 2 immunizations and in 100μl for the last immunization. No adjuvant was used. Sera were collected pre-immunization on day 1 and subsequently on days 22 and 36. The pre and post-immunization sera from these animals were evaluated by enzyme linked immunosorbent assay (ELISA) assay to assess for the development of antibodies directed against $IroN_{ec}$.

**b) ELISA assay for detection of antibodies directed against IroNec:**

ELISA assays were performed using Immulon 2 HB plates (DYNEX), coated with 75ng of purified IroNec/well. The serum dilution was 1/1000. The conjugate used was Peroxidase-Labeled Goat anti Mouse IgG + IgM at a concentration of 1/10,000. $IroN_{ec}$ was adsorbed to the plate overnight. In the morning, the plate was washed and blocked using PIERCE Superbloc. After blocking, the plate was washed again and incubated with the diluted sera for 2 hours. The plate was washed

again and incubated with the diluted conjugate for 1 hr before it was developed. Readings were measured at 450nm on an automated ELISA plate reader.

The results of this assay are summarized in Table 4. As can be seen, a significant increase in antibodies directed against $IroN_{ec}$ developed in mice immunized with $IroN_{ec}$ (Groups 2 and 3) when compared with non-immunized controls (Group 1). Further, every animal in Group 2 and 3 had an antibody response of a similar magnitude. However, there was no difference in the magnitude of antibody response between the animals immunized with a total of 70$\mu$g of $IroN_{ec}$ (Group 2) when compared to animals immunized with a total of 150 $\mu$g of $IroN_{ec}$ (Group 3). These results demonstrate that $IroN_{ec}$ is antigenic, a critical property of a vaccine candidate.

Table 4:
Immunization with $IroN_{ec}$ elicits a strong antibody response in mice

| Immunizing regimen | ELISA $OD_{600}$ mean±SEM | Post/Pre ratio |
|---|---|---|
| Controls | 0.043±0.0007 (pre)<br>0.090±0.010 (post) | 2.1 |
| Immunized with 70$\mu$g of $IroN_{ec}$ | 0.084±0.015 (pre)<br>1.31±0.022 (post) | 15.6 |
| Immunized with 150$\mu$g of $IroN_{ec}$ | 0.099±0.008 (pre)<br>1.38±0.015 (post) | 14.7 |

**Example 6**

This embodiment demonstrates that immunization with $IroN_{ec}$ protects mice against challenge with the *Escherichia coli* strain CP9. Having established that immunization with $IroN_{ec}$ results in antibody production

directed against $IroN_{ec}$, experiments were performed to evaluate whether this antibody response was protective against challenge with the homologous *E. coli* strain CP9. For initial studies an intra-peritoneal challenge model was used.

**a) Intraperitoneal challenge model:**

Both male and female mice (18-22 grams) underwent the immunizing regimen described above using purified $IroN_{ec}$ as the immunogen. The non-immunized control Group 1 consisted of 15 animals. Group 2 (N=12) and Group 3 (N=15) were immunized with a total of 70$\mu$g and 150$\mu$g of purified $IroN_{ec}$ respectively. Sera were obtained prior to and after immunization. Twelve days after the third and final immunizing dose of $IroN_{ec}$ was administered, animals underwent intra-peritoneal challenge with four different titers of the *E. coli* strain CP9. This infection model results in a systemic infection that may be lethal, depending on the magnitude of the challenge titer. For these experiments, challenge titers were utilized that would result in a 50-75% mortality rate in control animals. In this manner, a protective effect of immunization with $IroN_{ec}$ could be identified. The measured endpoints of this study were 1) death and 2) hepatic and splenic bacterial titers. Animals were observed post- bacterial challenge. Upon death, the liver and spleen were immediately removed, homogenized, and titered for bacterial counts via serial 10-fold dilutions. If an animal was still alive 18 hours after bacterial challenge it was sacrificed and bacterial titers of the liver and spleen were performed as described above. Since the antibody response against $IroN_{ec}$ was the same for Groups 2 and 3, these groups were pooled for the analysis presented below.

When antibody response was evaluated by ELISA, none

of the control animals possessed significant pre-existing antibody to $IroN_{ec}$, nor did a significant antibody response occur after sham immunization. In contrast, all of the animals immunized with $IroN_{ec}$ developed a significant antibody response against it. The results are presented in Table 5.

Table 5

Immunization with $IroN_{ec}$ decreases mortality in mice challenged intraperitoneally with live *Escherichia coli* strain CP9

| Immunizing regimen | no. dead/no. Injected (%) | $LD_{50}$[&] |
|---|---|---|
| Negative controls not immunized with $IroN_{ec}$ | | |
| $2.5 \times 10^6$ cfu[#] | 1/3 (33) | $3.83 \times 10^6$ cfu |
| $8.1 \times 10^6$ | 3/4 (75) | |
| $2.5 \times 10^7$ | 4/4 (100) | |
| $8.1 \times 10^7$ | 4/4 (100) | |
| Immunized animals (immunized with $IroN_{ec}$) | | |
| $2.5 \times 10^6$ | 1/6 (16.6) | $7.84 \times 10^6$ cfu |
| $8.1 \times 10^6$ | 3/6 (50) | |
| $2.5 \times 10^7$ | 5/7 (71.4) | |
| $8.1 \times 10^7$ | 7/7 (100) | |

& = bacterial challenge dose needed for death in 50% of mice

# = bacterial colony forming units

As shown in Table 5, immunization with $IroN_{ec}$ both resulted in a decreased mortality in immunized animals (higher $LD_{50}$ dose, $7.84 \times 10^6$ cfu ) when compared to unimmunized controls (lower $LD_{50}$ dose, $3.83 \times 10^6$ cfu). Further, immunized control animals had diminished growth of CP9 in the liver and spleen when compared to unimmunized control animals (Table 6).

Table 6

Immunization with $IroN_{ec}$ diminishes bacterial growth in the liver and spleen of mice challenged intraperitoneally with live *Escherichia coli* strain CP9

| Immunizing regimen | Hepatic & Splenic growth/no. Injected (%) | $GD_{50}$[&] |
|---|---|---|
| Negative controls (not immunized with $IroN_{ec}$) | | |
| $2.5 \times 10^6$ cfu[#] | 1/3 (33) | $2.58 \times 10^6$ cfu |
| $8.1 \times 10^6$ | 4/4 (75) | |
| $2.5 \times 10^7$ | 4/4 (100) | |
| $8.1 \times 10^7$ | 4/4 (100) | |
| Immunized animals (immunized with $IroN_{ec}$) | | |
| $2.5 \times 10^6$ | 1/6 (16.6) | $7.84 \times 10^6$ cfu |
| $8.1 \times 10^6$ | 3/6 (50) | |
| $2.5 \times 10^7$ | 5/7 (71.4) | |
| $8.1 \times 10^7$ | 7/7 (100) | |

& = bacterial challenge dose needed for growth of CP9 in the liver and spleen of 50% of the mice

# = bacterial colony forming units

In summary, immunization with purified $IroN_{ec}$ resulted in the development of antibodies directed against it. Further, these antibodies resulted in protection against subsequent bacterial challenge as shown by a diminished mortality rate and diminished growth of the model pathogen CP9 in liver and spleen.

From the foregoing, it will be obvious to those skilled in the art the various modifications in the above-described methods and compositions can be made without departing from the spirit and scope of the invention. Accordingly, the invention may be embodied in other specific forms without departing from the spirit or essential characteristics thereof. Present examples and embodiments, therefore, are to be considered in all respects as illustrative and not restrictive, and all changes which come within the meaning and range of equivalency of the specifications are therefore intended to be embraced therein.

What is claimed is:

1. An isolated polynucleotide comprising a polynucleotide sequence selected from the group consisting of:
(a) a polynucleotide having at least 90% homology to a polynucleotide encoding a polypeptide of SEQ ID NO:8 and
(b) a polynucleotide which is complementary to the polynucleotide of (a).

2. The isolated and purified polynucleotide of claim 1, wherein the polynucleotide in (a) has at least 95% homology to a polynucleotide encoding a polypeptide of SEQ ID NO:8.

3. The isolated and purified polynucleotide of claim 2, wherein the polynucleotide in (a) has at least 97% homology to a polynucleotide encoding a polypeptide of SEQ ID NO:8.

4. The isolated and purified polynucleotide of claim 3, wherein the polynucleotide in (a) is a polynucleotide encoding a polypeptide of SEQ ID NO:8.

5. The isolated and purified polynucleotide of claim 4, wherein the polynucleotide in (a) encoding a polypeptide of SEQ ID NO:8 has a sequence of SEQ ID NO:1.

6. A recombinant vector comprising the polynucleotide of claim 1.

7. A recombinant vector comprising the polynucleotide of claim 4.

8. A recombinant vector comprising a nucleotide sequence encoding one or more antigenic epitopes of $IroN_{ec}$.

9. A pure peptide of SEQ ID NO:8

10. A pure antigenic peptide, polypeptide or protein having one or more epitopes of $IroN_{ec}$, wherein $IroN_{ec}$ has the sequence of SEQ ID NO:8.

11. The peptide or protein of claim 9, wherein the peptide, polypeptide or protein is produced recombinantly.

12. An antigenic formulation comprising a pure peptide, polypeptide or protein having one or more epitopes of $IroN_{ec}$, wherein $IroN_{ec}$ has the sequence of SEQ ID NO:8.

13. The antigenic formulation of claim 13 further comprising a pharmaceutical carrier.

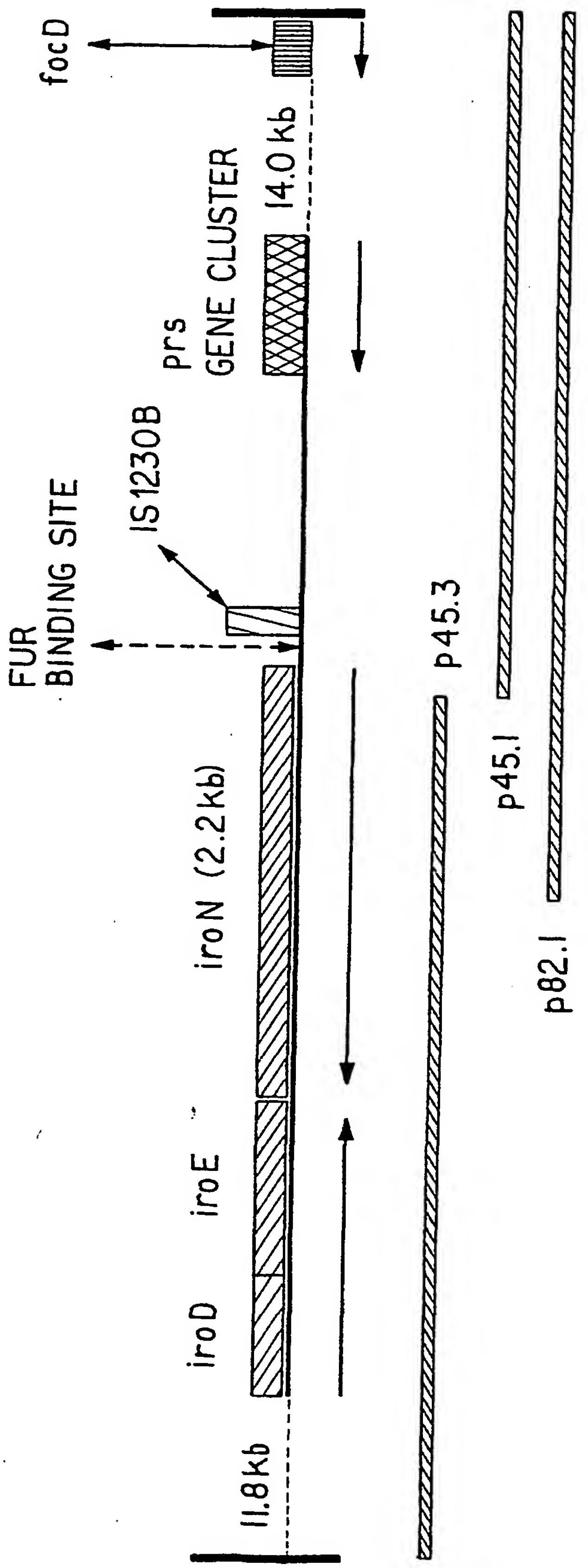
focD
14.0 kb
prs
GENE CLUSTER
IS1230B
FUR
BINDING SITE
iroN (2.2kb)
iroE
iroD
11.8kb
p45.3
p45.1
p82.1

FIG. 1

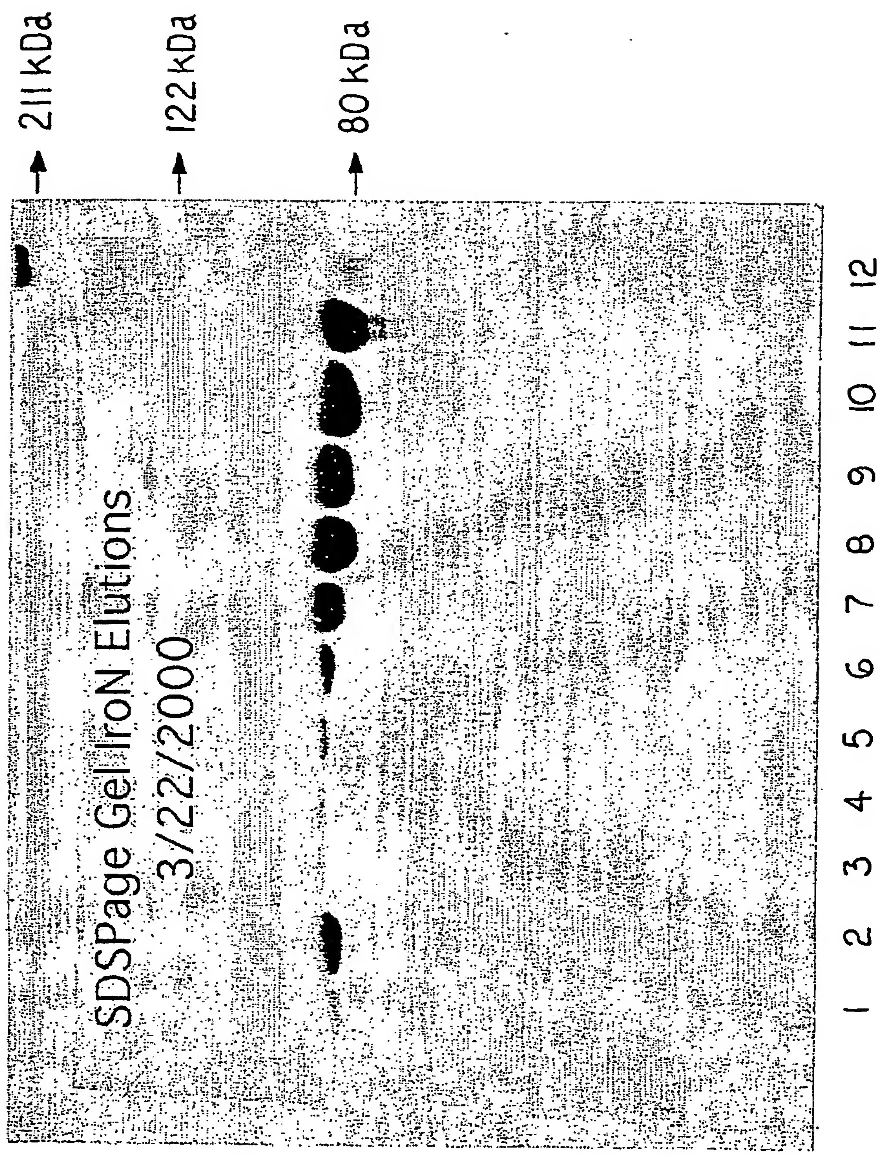
211 kDa
122 kDa
80 kDa
SDSPage Gel IroN Elutions
3/22/2000
1 2 3 4 5 6 7 8 9 10 11 12

FIG. 2A

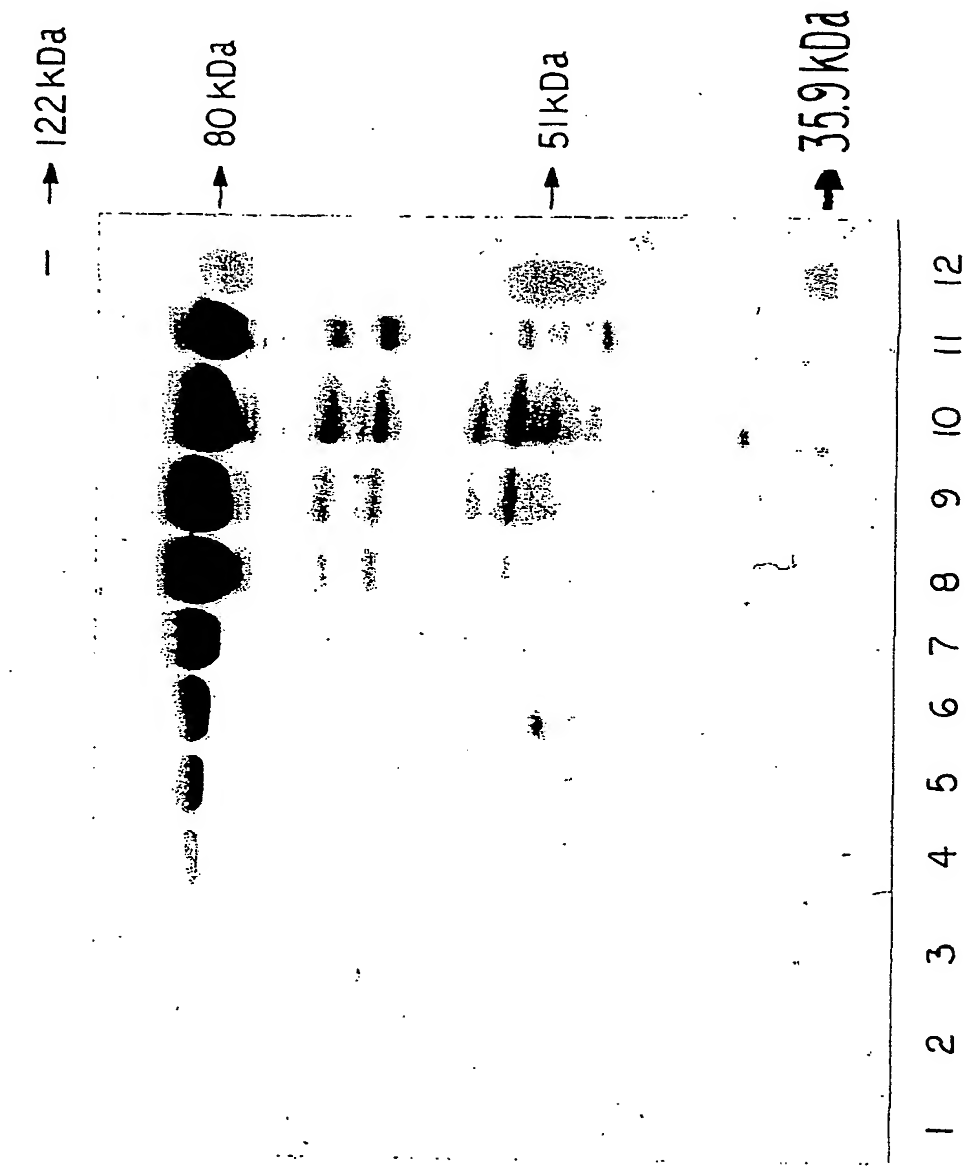


FIG. 2B

SEQUENCE LISTING

<110> Russo, Thomas A.

<120> Identification of A Vaccine Candidate from an Extraintestinal Strain of E. coli

<130> 11520.0214

<140> PCT/US00/26117
<141> 2000-09-22

<150> US 60/155,621
<151> 1999-09-22

<160> 10

<210> 1
<211> 2495
<212> DNA
<213> Escherichia coli

<220>

<400> 1
```
ctctgactcg cccgtcacga ctggaatcat gaaggcgttc aagttcagct    50
ttttgttgct cagtaatgaa tattttcatg ttggacagca tgatcttctg   100
ttgctgaaaa atcaagcatc ttcaatgatc acgggtatat atcaatttga   150
tatttatgaa aaatcaatag atttatcttg tgagggattg ctttttcttt   200
tgctaaatga tattaattat cattcttata taaaaacaat aatggcagta   250
agaaatatct ggcaaggatg tgagcttaac gatcaaacat tttaagttag   300
gcatttatta gggaatagga atgagaatta acaaaatcct ctggtcgcta   350
actgtgctcc tagttgggtt gaatagccag gtatcagtag ccaaatactc   400
cgacgatgat aatgacgaga ctctggtggt ggaagccacc gctgagcagg   450
tattaaaaca gcagccgggc gtgtcggtta ttaccagcga ggatattaaa   500
aagacccctc cggtaaacga cctttcagat attattcgta aaatgcctgg   550
tgttaatctt accggcaata gcgcctcggg cacacgcggt aataaccgcc   600
agatcgatat tcgtggtatg ggccggaaa acaccttaat tttaattgat   650
ggtgtaccgg tgacgtcacg taactccgtg cgttatagct ggcgtgggga   700
gcgtgatacc cgcggtgaca ccaactgggt gccaccggaa caggttgagc   750
gtattgaagt gatccgcggc cctgcggcgg cgcgctacgg ttcgggggcc   800
gccgggggg tggtgaacat cattaccaaa cgtcccacca acgactggca   850
```

cggttcgctg tcgttataca ccaaccagcc ggaaagtagc gaagagggcg 900
ctacgcgtcg cgccaatttc agccttagtg ggcctctggc tggtgatgct 950
cttaccacgc gtttgtatgg taacctgaat aaaacggatg ctgacagttg 1000
ggatattaat tctccggtcg gtacgaaaaa cgcagccggg catgaagggg 1050
tacgtaacaa agatattaac ggcgttgtct cgtggaaatt aaatccgcag 1100
cagattctcg atttcgaagt cggatatagc cgccagggga atatctatgc 1150
gggcgatacg cagaacagtt cttccagtgc agttaccgaa agcctggcaa 1200
aatccggcaa agagacgaac cgcctgtacc gacagaatta tggcattacg 1250
cataatggta tctgggactg gggacaaagt cgctttggtg tttattacga 1300
gaaaaccaat aatacccgca tgaatgaagg attatccggc ggtggtgaag 1350
gacgtatttt agcgggtgaa aagtttacga ccaatcgcct gagttcctgg 1400
cgaaccagcg gtgagcttaa tattcctttg aatgtgatgg ttgatcaaac 1450
gctgaccgtt ggtgcagagt ggaaccgcga taagctcgat gatccttcct 1500
ctaccagcct gacggtgaat gacagagata tcagcggtat ttctggctct 1550
gctgcggatc gcagcagtaa aaatcattct caaatcagtg cgctgtatat 1600
tgaagataac attgagccgg ttcctggcac gaatatcatt cccggcctgc 1650
gctttgatta tctcagcgac tccggcggga acttcagccc cagtctgaat 1700
ctttcgcagg aattgggcga ttatttcaaa gtcaaagcag gggttgcccg 1750
aacctttaaa gccccaaacc tgtatcaatc cagtgaaggc tatctgctct 1800
actcgaaagg caatggctgt ccaaaagata ttacatcagg cgggtgctac 1850
ctgatcggta ataaagatct cgatccggaa atcagcgtca ataaagaaat 1900
tggactggag ttcacctggg aagattacca cgcaagtgtg acctacttcc 1950
gcaatgatta ccagaataag atcgtggccg gggataacgt tatcgggcaa 2000
accgcttcag gcgcatatat cctcaagtgg cagaatggcg ggaaagctct 2050
ggtggacggt atcgaagcca gtatgtcttt cccactggtg aaagagcgtc 2100
tgaactggaa taccaatgcc acatggatga tcacttcgga gcaaaaagac 2150
accggtaatc ctctgtcggt catcccgaaa tatactatca ataactcgct 2200
taactggacc atcacccagg cgttttctgc cagcttcaac tggacgttat 2250
atggcagaca aaaaccgcgt actcatgcgg aaacccgcag tgaagatact 2300

ggcggtctgt caggtaaaga gctgggcgct tattcactgg tggggacgaa 2350

cttcaattac gatattaata aaaatctgcg tcttaatgtc ggcgtcagta 2400

atatcctcaa taaacagatc ttccgatctt ctgaaggggc gaatacctat 2450

aacgagccag gccgggctta ttatgccgga gttaccgcat cattc 2495

<210> 2
<211> 18
<212> DNA
<213> artificial sequence

<220>
<223> completely synthesized primer

<400> 2
gatcaagaga caggatga 18

<210> 3
<211> 18
<212> DNA
<213> artificial sequence

<220>
<223> completely synthesized primer

<400> 3
tgatcctcgc cgtactgc 18

<210> 4
<211> 15
<212> DNA
<213> artificial sequence

<220>
<223> completely synthesized primer

<400> 4
aatatcgccc tgagc 15

<210> 5
<211> 18
<212> DNA
<213> artificial sequence

<220>
<223> completely synthesized primer

<400> 5
catgttagga ggtcacat 18

<210> 6
<211> 18
<212> DNA
<213> artificial sequence

<220>
<223> completely synthesized primer

<400> 6
gacgccgaca ttaagacg 18

<210> 7
<211> 19
<212> DNA
<213> artificial sequence

<220>
<223> completely synthesized primer

<400> 7
aagtcaaagc aggggttgc 19

<210> 8
<211> 725
<212> PRT
<213> Escherichia coli

<220>
<400> 8
```
Met Arg Ile Asn Lys Ile Leu Trp Ser Leu Thr Val Leu Leu Val
                5                   10                  15

Gly Leu Asn Ser Gln Val Ser Val Ala Lys Tyr Ser Asp Asp Asp
                20                  25                  30

Asn Asp Glu Thr Leu Val Val Glu Ala Thr Ala Glu Gln Val Leu
                35                  40                  45

Lys Gln Gln Pro Gly Val Ser Val Ile Thr Ser Glu Asp Ile Lys
                50                  55                  60

Lys Thr Pro Pro Val Asn Asp Leu Ser Asp Ile Ile Arg Lys Met
                65                  70                  75

Pro Gly Val Asn Leu Thr Gly Asn Ser Ala Ser Gly Thr Arg Gly
                80                  85                  90

Asn Asn Arg Gln Ile Asp Ile Arg Gly Met Gly Pro Glu Asn Thr
                95                  100                 105

Leu Ile Leu Ile Asp Gly Val Pro Val Thr Ser Arg Asn Ser Val
                110                 115                 120

Tyr Ser Trp Arg Gly Glu Arg Asp Thr Arg Gly Asp Thr Asn Arg
                125                 130                 135

Trp Val Pro Pro Glu Gln Val Glu Arg Ile Glu Val Ile Arg Gly
                140                 145                 150

Pro Ala Ala Ala Arg Tyr Gly Ser Gly Ala Ala Gly Gly Val Val
                155                 160                 165

Asn Ile Ile Thr Lys Arg Pro Thr Asn Asp Trp His Gly Ser Leu
```

170 175 180

Ser Leu Tyr Thr Asn Gln Pro Glu Ser Ser Glu Glu Gly Ala Thr
185 190 195

Arg Arg Ala Asn Phe Ser Leu Ser Gly Pro Leu Ala Gly Asp Ala
200 205 210

Leu Thr Thr Arg Leu Tyr Gly Asn Leu Asn Lys Thr Asp Ala Asp
215 220 225

Ser Trp Asp Ile Asn Ser Pro Val Gly Thr Lys Asn Ala Ala Gly
230 235 240

His Glu Gly Val Arg Asn Lys Asp Ile Asn Gly Val Val Ser Trp
245 250 255

Lys Leu Asn Pro Gln Gln Ile Leu Asp Phe Glu Val Gly Tyr Ser
260 265 270

Arg Gln Gly Asn Ile Tyr Ala Gly Asp Thr Gln Asn Ser Ser Ser
275 280 285

Ser Ala Val Thr Glu Ser Leu Ala Lys Ser Gly Lys Glu Thr Asn
290 295 300

Arg Leu Tyr Arg Gln Asn Tyr Gly Ile Thr His Asn Gly Ile Trp
305 310 315

Asp Trp Gly Gln Ser Arg Phe Gly Val Tyr Tyr Glu Lys Thr Asn
320 325 330

Asn Thr Arg Met Asn Glu Gly Leu Ser Gly Gly Gly Glu Gly Arg
335 340 345

Ile Leu Ala Gly Glu Lys Phe Thr Thr Asn Arg Leu Ser Ser Trp
350 355 360

Arg Thr Ser Gly Glu Leu Asn Ile Pro Leu Asn Val Met Val Asp
365 370 375

Gln Thr Leu Thr Val Gly Ala Glu Trp Asn Arg Asp Lys Leu Asp
380 385 390

Asp Pro Ser Ser Thr Ser Leu Thr Val Asn Asp Arg Asp Ile Ser
395 400 405

Gly Ile Ser Gly Ser Ala Ala Asp Arg Ser Ser Lys Asn His Ser
410 415 420

Gln Ile Ser Ala Leu Tyr Ile Glu Asp Asn Ile Glu Pro Val Pro
425 430 435

Gly Thr Asn Ile Ile Pro Gly Leu Arg Phe Asp Tyr Leu Ser Asp
440 445 450

Ser Gly Gly Asn Phe Ser Pro Ser Leu Asn Leu Ser Gln Glu Leu
455 460 465

Gly Asp Tyr Phe Lys Val Lys Ala Gly Val Ala Arg Thr Phe Lys
470 475 480

Ala Pro Asn Leu Tyr Gln Ser Ser Glu Gly Tyr Leu Leu Tyr Ser
485 490 495

Lys Gly Asn Gly Cys Pro Lys Asp Ile Thr Ser Gly Gly Cys Tyr
500 505 510

Leu Ile Gly Asn Lys Asp Leu Asp Pro Glu Ile Ser Val Asn Lys
515 520 525

Glu Ile Gly Leu Glu Phe Thr Trp Glu Asp Tyr His Ala Ser Val
530 535 540

Thr Tyr Phe Arg Asn Asp Tyr Gln Asn Lys Ile Val Ala Gly Asp
545 550 555

Asn Val Ile Gly Gln Thr Ala Ser Gly Ala Tyr Ile Leu Lys Trp
560 565 570

Gln Asn Gly Gly Lys Ala Leu Val Asp Gly Ile Glu Ala Ser Met
575 580 585

Ser Phe Pro Leu Val Lys Glu Arg Leu Asn Trp Asn Thr Asn Ala
590 595 600

Thr Trp Met Ile Thr Ser Glu Gln Lys Asp Thr Gly Asn Pro Leu
605 610 615

Ser Val Ile Pro Lys Tyr Thr Ile Asn Asn Ser Leu Asn Trp Thr
620 625 630

Ile Thr Gln Ala Phe Ser Ala Ser Phe Asn Trp Thr Leu Tyr Gly
635 640 645

Arg Gln Lys Pro Arg Thr His Ala Glu Thr Arg Ser Glu Asp Thr
650 655 660

Gly Gly Leu Ser Gly Lys Glu Leu Gly Ala Tyr Ser Leu Val Gly
665 670 675

Thr Asn Phe Asn Tyr Asp Ile Asn Lys Asn Leu Arg Leu Asn Val
680 685 690

Gly Val Ser Asn Ile Leu Asn Lys Gln Ile Phe Arg Ser Ser Glu
695 700 705

Gly Ala Asn Thr Tyr Asn Glu Pro Gly Arg Ala Tyr Tyr Ala Gly
710 715 720

Val Thr Ala Ser Phe
725

<210> 9
<211> 33
<212> DNA
<213> Artificial Sequence

<220>
<223> forward primer for PCR

<400> 9
cgcgcgcgga tccgacgaga ctctggtggt gga 33

<210> 10
<211> 33
<212> DNA
<213> Artificial Sequence

<220>
<223> reverse primer for PCR

<400> 10
cgcgcgcaag cttgaatgat gcggtaactc cgg 33

# INTERNATIONAL SEARCH REPORT

International application No.
PCT/US00/26117

**A. CLASSIFICATION OF SUBJECT MATTER**

IPC(7) :Please See Extra Sheet.
US CL :536/23.1, 23.7; 435/69.1, 69.3, 71.1, 320.1; 530/350; 424/250.1
According to International Patent Classification (IPC) or to both national classification and IPC

**B. FIELDS SEARCHED**

Minimum documentation searched (classification system followed by classification symbols)
U.S. : 536/23.1, 23.7; 435/69.1, 69.3, 71.1, 320.1; 530/350; 424/250.1

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)
WEST, DIALOG, MEDLINE, EMBASE, BIOSIS, CA SEARCH

**C. DOCUMENTS CONSIDERED TO BE RELEVANT**

| Category* | Citation of document, with indication, where appropriate, of the relevant passages | Relevant to claim No. |
|---|---|---|
| X, P<br>----<br>Y, P | JOHNSON et al. Molecular Epidemiological and Phylogenetic Associations of Two Novel Putative Virulence Genes, iha and ironE.coli' among Escherichia coli Isolates from Patients with Urosepsis. Infect. Immun. May 2000. Vol. 68. No. 5. pages 3040-3047, see entire document. | 1-8<br>----<br>9-13 |
| A | RUSSO et al. Identification, Genomic Organization, and Anaylsis of the Group III Capsular Polysaccharide Genes kpsD, kpsM, kpsT, and kpsE from an Extrointestinal Isolate of Escherichia coli (CP9, O4/K54/H5). J. Bacteriol. January 1998, Vol. 180. No. 2. pages 338-349, see entire document. | 1-13 |

[X] Further documents are listed in the continuation of Box C. [ ] See patent family annex.

* Special categories of cited documents:

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"&" document member of the same patent family

Date of the actual completion of the international search
30 NOVEMBER 2000

Date of mailing of the international search report
02 FEB 2001

Name and mailing address of the ISA/US
Commissioner of Patents and Trademarks
Box PCT
Washington, D.C. 20231
Facsimile No. (703) 305-3230

Authorized officer
JENNIFER GRASER
TERRY J. DEY
PARALEGAL SPECIALIST
TECHNOLOGY CENTER 1600
Telephone No. (703) 308-0196

## INTERNATIONAL SEARCH REPORT

Inte l application No.
PCT/US00/26117

A. CLASSIFICATION OF SUBJECT MATTER:
IPC (7):

C07H 21/02, 21/04; C12P 21/06, 21/04; C12N 15/09, 15/00; C07K 1/00; A61K 39/095

Form PCT/ISA/210 (extra sheet) (July 1998) ★

C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

| Category* | Citation of document, with indication, where appropriate, of the relevant passages | Relevant to claim No. |
|---|---|---|
| A, P | RUSSO et al. Identification of Genes in an Extraintestinal Isolate of Escherichia coli with Increased Expression after Exposure to Human Urine. Infect. Immun. October 1999, Vol. 67. No. 10. pages 5306-5314, see entire document. | 1-13 |